Naval Research Laboratory

Washington, DC 20375-5320

AD-A255 636

NRL/MR/5521—92-7107

(2)

Multi-Access Strategies for Voice/Data Integration in Heterogeneous Mixed-Media Packet Radio Networks

EVAGGELOS GERANIOTIS

Locus, Inc.
Alexandria, Virginia

and

University of Maryland
College Park, Maryland

and

MOHSEN SOROUSHNEJAD

University of Maryland
College Park, Maryland

Prepared for

Communication Systems Branch
Information Technology Division

September 8, 1992

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 8, 1992	Interim Report 3/91-11/91

4. TITLE AND SUBTITLE

Multi-Access Strategies for Voice/Data Integration in Heterogeneous Mixed-Media Packet Radio Networks

5. FUNDING NUMBERS

PE - 61153N
PR - RR015-09-41
WU - DN159-036

6. AUTHOR(S)

Evaggelos Geraniotis and Mohsen Soroushnejad

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/5521—92-7107

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Part of this research was performed on site at the Naval Research Laboratory. E. Geraniotis is with the University of Maryland and LOCUS Inc. M. Soroushnejad is with the University of Maryland.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The problem of voice/data integration in a mixed-media packet radio network employing both ground radio links and SATCOM radio links is investigated. Multi-access strategies are devised to take advantage of different characteristics of both radio channels to better serve the requirements of both traffic types. Code-division multiple-access with movable boundary in the code domain is used in the ground subnetwork to serve the voice traffic and the retransmitted data traffic while framed ALOHA with movable boundary is used on the satellite subnetwork for both data and voice.

Based on a Markovian model of the system, a complete analysis of the channel access protocols for both traffic types is given. The performance of the multiple-access schemes used in this mixed-media network is evaluated in terms of voice and data throughput, voice blocking probability, and data delay. It is observed that splitting of retransmission traffic between the satellite and ground subnet may increase the overall data throughput when the voice load is high.

14. SUBJECT TERMS

Communications network
Voice/data integration
Multiple access

15. NUMBER OF PAGES 56

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

CONTENTS

MULTI-ACCESS STRATEGIES FOR VOICE/DATA INTEGRATION IN HETEROGENEOUS MIXED-MEDIA PACKET RADIO NETWORKS

1. INTRODUCTION

In a variety of commercial and military communication networks, it is possible that the network links employ different communication media with markedly different characteristics. The different characteristics could be different frequency bands, data rates (channel capacities), error rates, delays, and network connectivity. We call these networks **heterogeneous mixed-media networks**, while we use the terminology of **multi-media networks** to refer to networks supporting different types of traffic, such as voice and data. The networks studied in this report actually fall into both of these classes.

As examples of such **heterogeneous mixed-media networks** consider radio networks, which use the HF (both groundwave and skywave), UHF, and EHF SATCOM frequency bands. In these networks, some of the links are terrestrial links (not necessarily line-of-sight), while others are satellite links. The scenario for the ground component of such networks may be of terrestrial based packet radio integrated networks that support mobile terminals. They can also be a subnet of a larger system of interconnected integrated networks that provide radio interface for some users of a backbone packet-switched network. The satellite component may involve a single satellite over the geographical area of the terrestrial subnetwork or even a network of multiple satellites covering a much broader geographical area with many distinct terrestrial networks.

Mixed-media networks are important, because, if suitable multiple-access and routing algorithms are devised, the overall network efficiency can be considerably increased by using all available resources (media, e.g., communication links over different frequency bands). Furthermore, for military applications and under stressed conditions, the combined use of all media may provide connectivity, whereas use of a single medium may not.

In the paper by D. Huynh, H. Kobayashi, and F. F. Kuo ([1]), a mixed-media network was considered consisting of a ground subnet and a satellite subnet. The ground subnet was connected with wired lines and used standard store-and-forward packet switching. The satellite subnet used contention-based channel access (ALOHA). The retransmissions of

Manuscript approved May 12, 1992

the packets that collided in the satellite channel were split between the ground links and the satellite link. Only data packets were transmitted through the network of [1].

The other issue of interest to this report is **multi-media integration** over a common channel. This is motivated by evolution of communication networks toward integrated services digital networks (ISDN) to accommodate random demands for service from a population of heterogeneous users with a variety of data types (such as interactive data, digital voice, video). These traffic types (in this report we restrict attention to voice and data) have different characteristics; voice can tolerate higher error rates but requires constant delays with no variability, whereas data can tolerate longer delays but require lower error rates.

Integration of different traffic types over a common channel requires a method to determine how the users should coordinate their transmissions to avoid the destructive interference which occurs when the traffic load over the channel increases. In other words, a channel-access scheme has to be followed by the terminals of possibly different traffic types in order to make efficient use of the channel. In a multi-media environment hybrids of circuit switching and packet switching techniques have been recommended (see [2]-[4]) which accommodate both the need of voice traffic for real-time delivery and the required low packet error probabilities of data traffic.

In the recent work of [5] fixed- and movable-boundary channel-access schemes for voice/data integration in wireless networks were introduced and their performance analyzed. Of particular interest were the movable-boundary schemes of [5], which allocate in a dynamic manner time slots to voice or data traffic according to the overall network traffic conditions. Moreover, in our recent work of [6] multiple-access schemes for voice/data integration in spread-spectrum packet radio networks were introduced and their performance evaluated. The introduction of movable-boundary schemes in the code domain was the gist of [6]. The work of both [5] and [6] pertained to networks with a single medium.

In this report, we consider a packet radio network with ground subnet using radio links and a satellite subnet. For example, HF links may be used for the ground subnet and EHF SATCOM links for the satellite subnet. Both voice and data packets are transmitted through the network. The terrestrial (or ground) channel is characterized by a bandwidth

W_1, delay τ_1, and packet error rate $P_{e,1}$. The corresponding features of the satellite (SATCOM) channel are W_2, τ_2, and $P_{e,2}$, respectively. It is assumed that $\tau_1 < \tau_2$ and $P_{e,1} > P_{e,2}$. These comparisons are typical of the relationships of the two channels in practical situations.

In this report efficient multiple-access schemes are devised which take advantage of the characteristics of the terrestrial and satellite channels above and accommodate the requirements of the different traffic types (voice and data). In accomplishing this, our report extends the work of [1] for purely data mixed-media networks to radio mixed-media networks with two traffic types (voice and data), our work of [6] for single-medium networks to mixed-media networks, and, to a certain extent the work of [5] for single-medium networks to the mixed-media case.

In the terrestrial radio network Code-Division Multiple-Access (CDMA) is used. This is motivated by the fact that CDMA can support a good number of asynchronous (at the frequency-hop or dwell-time level) simultaneous users, it mitigates the effects of fading and narrowband interference, it combats jamming, and it provides selective addressing/reception capability. Moreover, the performance of CDMA schemes degrades gracefully as the interference level (from any source) increases; this can be used efficiently for voice traffic given the graceful degradation of voice quality with increasing in-band (or co-channel) interference. Similar feedback and voice/data integration schemes as those developed in [6] (termed schemes A, B, and C in that report) can be used for the ground channel.

On the other hand, a framed protocol employing reservations for the voice traffic and ALOHA with retransmission control for the data traffic is used by the satellite channel; this is a version of the framed ALOHA protocol (see [7]-[8]), modified to accommodate both voice and data traffic.

The choice of channels and basic transmission policies for the two traffic types is described and justified in the following. The voice traffic uses primarily the terrestrial subnetwork and, if necessary, the satellite subnetwork as well. This is motivated by the requirement for short delay times and low variability of delay of the voice calls. Also a higher level of interference is present in the terrestrial channel making it less reliable.

However, since voice can tolerate higher levels of interference than data, the selection of the terrestrial network as the prime carrier of voice traffic is well justified.

Due to the more stringent requirements for smaller packet error probability and the tolerance of longer delays of the data users, data traffic is first transmitted through the more reliable satellite channel and, in the event of packet collisions, some of the retransmitted data packets are sent through the terrestrial subnetwork, while the remaining packets are sent through the satellite channel. This splitting is performed dynamically based on the current state of voice and data traffic in the two channels.

The mixed-media network of interest to this report is illustrated in Figure 1 with examples of different types of links.

In our report we study the tradeoffs among the following performance measures: throughput of data and voice packets, delay of data packets, and voice call blocking probability. The analysis is in principle applicable to several different strategies for upgrading the retransmission probabilities and the splitting probabilities; however, for the sake of clarity in the presentation and simplicity in generation of the numerical results only one such strategy is analyzed and and evaluated.

This report is organized as follows. Section 2 presents the system model and the detailed protocol operation for the ground and satellite subnetworks. In Section 3, we develop a Markovian model for voice traffic analysis. Section 4 follows with the model and the analysis for data traffic. In Section 5 numerical results about the system performance are presented.

2. SYSTEM MODEL

We consider a packet radio network with a population of M_d data users and M_v voice users. The network consist of a ground subnet (G-subnet), using ground radio links, and a satellite subnet (S-subnet). The time axis of the satellite channel is grouped into frames of length L slots each. The frame length should be at least as large as the round trip propagation delay of the satellite link; this is elaborated upon in Section 2.1 below. Time over the ground channel is also measured in frames but there is only one time slot in each frame. The basic time unit is the slot which represents the time allowed for the

transmission of a single data or voice packet. The frames of the ground channel and the frames of the satellite channel are perfectly aligned. It is assumed that the bandwidth of the satellite channel is equal to L times the bandwidth of the ground channel for the information signals. Therefore, the information rate over the satellite channel is L times the information rate over the ground channel and, since each user transmitting over the satellite channel uses only one slot in a frame, the same amount of information per user is transmitted during one frame of the satellite channel and one frame of the ground channel. Figures 2 and 3 illustrate the protocol operation over the satellite and ground subnets, respectively, over three consecutive frames. The detailed operation is described below.

2.1. Channel Access for Voice Traffic

Code-division multiple-accessing (CDMA) is used over the ground subnet to accommodate several voice calls simultaneously. In our model, packet transmissions start at common clock instances and packets have constant length of N symbols. Each user (data or voice node) employs a random frequency hopping pattern for the transmission of its packets. The frequency spectrum is divided into q frequency bins and each symbol is transmitted at a frequency chosen from the q frequencies with equal probability, independently of the frequencies chosen for other symbols. We assume that some form of Forward Error Control (FEC) coding is employed, for which up to ν symbol errors can be corrected. We also assume that a packet consists of exactly one codeword with N code symbols. Therefore, a packet is declared successfully transmitted, if at most ν symbol errors occur. In particular, if Reed-Solomon (N, K) codes with erasures-correction decoding are employed, $\nu = N - K$ erasures can be corrected. Typically, M-ary FSK (Frequency-Shift-Keying) data modulation with $M = N$ and noncoherent demodulation is employed in frequency-hopped spread-spectrum (FH/SS) systems. K M-ary symbols are encoded on N Reed-Solomon code symbols and each code symbol is transmitted on a single frequency hop. Each voice packet of a typical user occupies exactly one frame over the ground channel.

We define the Multiple Access Capability (MAC) index K_v^g as the number of voice users that can be accommodated simultaneously such that the expected packet error prob-

ability of voice traffic remains below a specified threshold. Specifically, the MAC indices K_v^g and K_d^g (for the data) are defined by

$$P_E(k) \leq P_E^v \quad \forall\, k \leq K_v^g \tag{2.1}$$

$$P_E(k) \leq P_E^d \quad \forall\, k \leq K_d^g \tag{2.2}$$

where P_E^v and P_E^d are the maximum tolerable voice and data packet error probabilities, respectively, and $P_E(k)$ is the packet error probability in the presence of k simultaneous packet transmissions, where k includes both voice and data users. In practice, $P_E^v > P_E^d$ and therefore $K_v^g > K_d^g$. If the total number of simultaneous users is $k \leq K_d^g$, then all k data or voice packets are received with acceptable error probability; if $K_d^g < k \leq K_v^g$, then among the k packets the voice packets are received with acceptable error probability, whereas the data packets have higher than the acceptable error probability; finally, if $K_v^g < k$, then all voice or data packets are received with unacceptable error probabilities. The model described above is referred to as the **threshold model** for other-user interference in CDMA systems.

In our analysis of data traffic over the ground subnet we also use the **graceful degradation model** for CDMA multiple-reception of data packets. According to this model, there is a non-zero probability of correct reception for any arbitrary number of packets which depends on the total number of simultaneously transmitted packets (data and voice) over the G-subnet. This model is described in detail in Section 4.2.

Notice that this discussion assumes the same code rate for both voice and data packets. It could be extended to systems with different code rates for the data and voice packets.

Up to K_v^s number of slots per frame may be allocated to the voice traffic on a reservation basis over the satellite subnet. One voice packet of a typical user occupies exactly one slot per frame over the satellite channel. The first slot of each frame over the satellite subnet is used by the voice nodes to reserve a slot for packet transmission in the next frame. We assume that the transponder employed for the satellite subnet is located on a geosynchronous satellite which results in a round-trip propagation delay of 0.27 seconds. Therefore, the frame length must be as large as the round-trip propagation delay, so that

the reservation information will be available to the users before the start of the next frame. The round-trip propagation delay is of the order of 10-13 slots for typical satellite networks.

Each idle voice node becomes active in a frame with probability P_v and attempts transmission at the beginning of the next frame. Each voice call generates a random number of packets geometrically distributed with parameter p. This means, that the probability that a voice call lasts for l frames (transmitting one packet per slot per frame over the satellite channel or one packet per frame over the ground channel) is

$$P\{\text{voice call duration} = l \text{ packets}\} = (1-p)^{l-1}p \quad \text{for } l = 1, 2, \cdots \tag{2.3}$$

The value $l = 0$ was excluded from (2.3), because it is assumed that a call that just became active can not terminate before it lasts for at least one packet (frame). Thus, the average duration of a voice call is

$$\sum_{l=1}^{\infty} l(1-p)^{l-1}p = \frac{1}{p} > 1 \tag{2.4}$$

which is larger than one packet.

It is assumed that some information about the number of established voice calls in progress is available to the nodes of the network at the beginning of each frame. This information may be obtained by observing (sensing) the channel activities during each frame. Several sensing schemes are possible that provide different amounts of information about the state of the channel.

Among the many possibilities, three different feedback and control schemes were considered in [5] as the most representative and meaningful. Scheme A characterized the situation where the nodes of the network observe the channel activities in each slot and obtain some feedback information based on the number of combined active voice and data terminals. In Scheme B, it was assumed that a central controller provides the nodes of the network with the exact number of established voice calls and data nodes in backlog. Scheme C is a variation of Scheme B, where the voice terminals do not use the information about the data backlog to regulate their access to the channel. Variations of Schemes A, B, and C were also considered according to which data and voice share the multiple-access

capability of the CDMA channel in a movable-boundary fashion with varying degrees of fairness in the initial and time-varying allocation of the MAC between them.

In this report we assume that channel sensing (feedback) makes available (i) the number of ongoing voice calls over both the G-subnet and the S-subnet to all voice users, and (ii) the data backlog over the G-subnet, the number of colliding data packets over the S-subnet, and the number of ongoing voice calls over the S-subnet and the G-subnet to all data users. The voice users do not use information about the data traffic. This is a variation of Scheme C of [5] modified to be suitable in mixed-media (ground/satellite networks) environment.

Voice-call blocking is used as a mechanism to control the flow of voice traffic. At the beginning of each frame, the voice nodes attempting to establish new calls observe the state information of the ground channel. If the number of established voice calls over the ground subnet is equal to the MAC index K_v^g, the voice nodes observe the state information of the satellite channel. If the number of established calls over the satellite channel equals the maximum capacity allocated to the voice traffic over the satellite subnet (i.e., K_v^s), then the new calls will be blocked and cleared. Otherwise, up to K_v^s number of calls may be established. With this access protocol for voice traffic, once a voice call is established, the voice packets either experience no delays (i.e., voice calls over the ground subnet), or have a constant delay of L slots (i.e., voice calls over the satellite subnet).

2.2. Channel Access for Data Traffic

Due to the high reliability of the satellite channel, the data traffic primarily goes through the satellite subnet. Actually, all data nodes initially transmit over the satellite channel. At the beginning of each frame, the data nodes have information about the number of established voice calls in that frame. This information determines the number of slots available for data packet transmission in that frame. We assume that each data terminal generates a new packet (independently of previous packets and other terminals) with probability P_d in each frame and remains idle with probability $1 - P_d$. It is thus possible that a data node generates a new packet in a frame although some of its previously generated packets have not yet been transmitted successfully. It is assumed that all

data nodes have at their disposal infinite size buffers (moderately large size buffers will suffice in most practical situations) where they store the backlogged packets waiting for retransmission and the newly arrived packets waiting for their first transmission.

The framed ALOHA protocol is used by the data nodes to access the satellite channel [7]-[8]. All the new data packet arrivals during a frame will be transmitted during that frame. The result of transmissions of data packets will be known by the end of the next frame. The data terminals that were involved in a collision event may retransmit the collided packets either through the satellite subnet or the ground subnet. The splitting of the retransmission traffic between the satellite and ground subnet may be controlled dynamically at the beginning of each frame. The data packets that join the backlog of data packets over the ground subnet are transmitted employing an ALOHA scheme with retransmission control. Since CDMA is used over the G-subnet that can support a maximum of K_d^g simultaneous data and voice packet transmissions at acceptable packet error probabilities, the retransmission probabilities for each frame depend on K_d^g and the number of ongoing voice calls over the ground network during this frame.

3. VOICE TRAFFIC ANALYSIS

We begin with the introduction of random processes that model the voice traffic. Let

N_i^{vg} = Number of voice calls in progress in the G-subnet at the beginning of frame i

N_i^{vs} = Number of voice calls in progress in the S-subnet at the beginning of frame i

A_i^v = Number of new voice call arrivals in frame i

E_i^g = Number of calls accepted by the G-subnet in frame i and transmitted in frame $i+1$

X_i^g = Number of calls terminated in frame i over the G-subnet

E_i^s = Number of calls accepted by the S-subnet in frame i and transmitted in frame $i+1$

X_i^s = Number of calls terminated in frame i over the S-subnet

Figure 4 depicts all components of voice traffic over the ground and satellite subnets.

The evolution of $\{N_i^{vg}\}$ and $\{N_i^{vs}\}$ is described by

$$\begin{aligned} N_{i+1}^{vg} &= N_i^{vg} + E_i^g - X_i^g \\ N_{i+1}^{vs} &= N_i^{vs} + E_i^s - X_i^s \end{aligned} \tag{3.1}$$

where

$$E_i^g = \begin{cases} 0 & N_i^{vg} \geq K_v^g \\ A_{i-1}^v & A_{i-1}^v \leq K_v^g - N_i^{vg}, \;\; N_i^{vg} < K_v^g \\ K_v^g - N_i^{vg} & A_{i-1}^v > K_v^g - N_i^{vg}, \;\; N_i^{vg} < K_v^g \end{cases} \tag{3.2a}$$

$$E_i^s = \begin{cases} A_{i-1}^v & A_{i-1}^v \leq K_v^s - N_i^{vs}, \;\; N_i^{vg} \geq K_v^g \\ K_v^s - N_i^{vs} & A_{i-1}^v > K_v^s - N_i^{vs}, \;\; N_i^{vg} \geq K_v^g \\ 0 & A_{i-1}^v \leq K_v^g - N_i^{vg}, \;\; N_i^{vg} < K_v^g \\ A_{i-1}^v - (K_v^g - N_i^{vg}) & 0 < A_{i-1}^v - (K_v^g - N_i^{vg}) \leq K_v^s - N_i^{vs}, \;\; N_i^{vg} < K_v^g \\ K_v^s - N_i^{vs} & A_{i-1}^v - (K_v^g - N_i^{vg}) > K_v^s - N_i^{vs}, \;\; N_i^{vg} < K_v^g \end{cases} \tag{3.3a}$$

or equivalently,

$$E_i^g = \begin{cases} 0 & N_i^{vg} \geq K_v^g \\ \min(A_{i-1}^v, K_v^g - N_i^{vg}) & N_i^{vg} < K_v^g \end{cases} \tag{3.2b}$$

$$E_i^s = \begin{cases} \min(A_{i-1}^v, K_v^s - N_i^{vs}) & N_i^{vg} \geq K_v^g \\ \min(A_{i-1}^v - E_i^g, K_v^s - N_i^{vs}) & N_i^{vg} < K_v^g \end{cases} \tag{3.3b}$$

Let $\underline{N}_i^v \stackrel{\Delta}{=} (N_i^{vs}, N_i^{vg})$. From the above, we see that N_i^{vs} and N_i^{vg} are dependent on each other, but the joint statistics of $\underline{N}_i^v$ can be shown to satisfy the Markovian property. The transition probabilities of $\underline{N}_i^v$ denoted by $P[\underline{N}_{i+1}^v = (n, m) \mid \underline{N}_i^v = (k, j)]$ may be

obtained by observing that

$$P[\underline{N}^v_{i+1} = (n,m) \mid \underline{N}^v_i = (k,j)]$$

$$= P[N^{vs}_{i+1} = n,\ N^{vg}_{i+1} = m \mid N^{vs}_i = k, N^{vg}_i = j]$$

$$= \sum_{l_1=\max\{0,m-j\}}^{\min\{M_v-k-j,K^g_v-j,m\}} \ \sum_{l_2=max\{0,n-k\}}^{\min\{M_v-k-j-l_1,K^s_v-k,n\}} P[\ E^g_i = l_1\ ,\ E^s_i = l_2 \mid N^{vs}_i = k, N^{vg}_i = j]$$

$$\cdot P[N^{vs}_{i+1} = n, N^{vg}_{i+1} = m | E^g_i = l_1, E^s_i = l_2, N^{vs}_i = k, N^{vg}_i = j] \tag{3.4}$$

The first term in the sum of (3.4) is given by

$$P[N^{vs}_{i+1} = n,\ N^{vg}_{i+1} = m \mid E^g_i = l_1\ ,\ E^s_i = l_2\ ,\ N^{vs}_i = k, N^{vg}_i = j]$$
$$= P[N^{vg}_{i+1} = m \mid E^g_i = l_1,\ N^{vg}_i = j] \ \cdot\ P[N^{vs}_{i+1} = n \mid E^s_i = l_2,\ N^{vs}_i = k] \tag{3.5}$$

Equation (3.5) holds because of the fact that when conditioned on ($E^g_i = l_1$, $E^s_i = l_2$), the number of accepted calls over the G-subnet and the S-subnet, and on ($N^{vs}_i = k, N^{vg}_i = j$), the number of ongoing calls at the beginning of frame i, N^{vs}_{i+1} and N^{vg}_{i+1}, (the number of ongoing calls at the beginning of frame $i+1$) are independent. Recall that memoryless geometric distributions determine the duration of each call.

The two terms in the right hand side of (3.5) are given by

$$P[N^{vg}_{i+1} = m \mid E^g_i = l_1,\ N^{vg}_i = j] = b(j, m - l_1, 1 - p) \tag{3.6}$$

and

$$P[N^{vs}_{i+1} = n \mid E^s_i = l_2,\ N^{vs}_i = k] = b(k, n - l_2, 1 - p). \tag{3.7}$$

where the binomial distribution is defined as

$$b(N, n, p) \triangleq \binom{N}{n} p^n (1-p)^{N-n} \quad ;0 \le n \le N. \tag{3.8}$$

In (3.6) and (3.7) $1 - p$ is the probability that an active call in the S-subnet during frame i continues during frame $i + 1$. The differences $m - l_1$ and $n - l_2$ are involved in

(3.6) and (3.7), respectively, because the newly accepted calls E_i^g and E_i^s over the G- and S-subnets during frame i are actually transmitted in frame $i+1$ and cannot terminate before transmitting at least one packet according to the model of (2.3).

The second term in the sum of (3.4) can be obtained from equations (3.2a) and (3.3a). As seen from (3.2a) and (3.3a) the two numbers of accepted calls E_i^g and E_i^s depend not only on the number of voice calls in progress N_i^{vg} and N_i^{vs} in frame i but through the number of new arrivals A_{i-1}^v they depend also on the number of voice calls in progress N_{i-1}^{vg} and N_{i-1}^{vs} in frame $i-1$. To capture this dependence accurately would transform the Markov chain of (3.4) from a first-order to a second-order one. This will greatly complicate the analysis. Instead, we make the following well justified approximation. Since the numbers of voice calls in progress N_i^{vg} and N_i^{vs} change very slowly from frame to frame, we can substitute N_i^{vg} and N_i^{vs} in (3.2a) and (3.3a) by N_{i-1}^{vg} and N_{i-1}^{vs}, respectively, without substantial loss in accuracy. This results in

$$P[\,E_i^g = l_1\,,\ E_i^s = l_2 \mid N_i^{vs} = k, N_i^{vg} = j] \approx P[\,E_i^g = l_1\,,\ E_i^s = l_2 \mid N_{i-1}^{vs} = k, N_{i-1}^{vg} = j]$$

$$= \begin{cases} b(M_v - k - j, l_1, P_v) & l_1 < K_v^g - j, l_2 = 0 \\ b(M_v - k - j, l_1 + l_2, P_v) & l_1 = K_v^g - j, l_2 < K_v^s - k \\ \sum_{\beta = l_1 + l_2}^{M_v - k - j} b(M_v - k - j, \beta, P_v) & l_1 = K_v^g - j, l_2 = K_v^s - k \\ 0 & elsewhere \end{cases} \tag{3.9}$$

Once the transition probability matrix of $\underline{N}_i^v$ is obtained, we can evaluate the stationary distribution of the process $\{\Pi^v(k,j)\}$, where $0 \le k \le K_v^s$ and $0 \le j \le \min\{K_v^g, M_v - k\}$; if we assume that $M_v \ge K_v^g + K_v^s$, then $0 \le j \le K_v^g$. The voice throughput is defined as the expected number of voice calls in progress per frame, and is given by

$$\eta_v \stackrel{\triangle}{=} \lim_{i \to \infty} E\{N_i^{vs} + N_i^{vg}\} = \sum_{k=0}^{K_v^s} \sum_{j=0}^{K_v^g} (k+j)\Pi^v(k,j) \tag{3.10}$$

Moreover, the voice-call blocking probability is given by

$$P_B = \frac{E\{\text{number of blocked calls per frame}\}}{E\{\text{number of arriving calls per frame}\}}$$

$$= \frac{\lim_{i\to\infty}\sum_{n=0}^{K_v^g+K_v^s}\sum_{l=K_v^g+K_v^s-n+1}^{M_v-n}[l-(K_v^g+K_v^s-n)]P[A_{i-1}^v=l|N_i^{vs}+N_i^{vg}=n]P[N_i^{vs}+N_i^{vg}=n]}{\lim_{i\to\infty}\sum_{n=0}^{K_v^g+K_v^s}\sum_{l=1}^{M_v-n}lP[A_{i-1}^v=l|N_i^{vs}+N_i^{vg}=n]P[N_i^{vs}+N_i^{vg}=n]}$$

$$= \frac{\sum_{n=0}^{K_v^g+K_v^s}\sum_{l=K_v^g+K_v^s-n+1}^{M_v-n}[l-(K_v^g+K_v^s-n)]b(M_v-n,l,P_v)\lim_{i\to\infty}P[N_i^{vs}+N_i^{vg}=n]}{\sum_{n=0}^{K_v^g+K_v^s}(M_v-n)P_v\lim_{i\to\infty}P[N_i^{vs}+N_i^{vg}=n]} \tag{3.11}$$

The above equation holds under the ergodicity assumption for the Markov Chain (N_i^{vs}, N_i^{vg}), so that the time averages involved in the practical definition of blocking probability are equivalent to the ensemble averages used in (3.11). Also we used the well justified approximation

$$P[A_{i-1}^v = l|N_i^{vs} + N_i^{vg} = n] \approx P[A_{i-1}^v = l|N_{i-1}^{vs} + N_{i-1}^{vg} = n] = b(M_v - n, l, P_v)$$

according to the discussion preceding (3.9). $K_v^g + K_v^s$ represents the maximum number of voice calls that can be accommodated by the ground and satellite subnets (total voice capacity of the system). According to (3.11), blocking occurs when the number of new (arriving) voice calls $A_{i-1}^v = l$ is larger than the system's voice capacity minus the number of current (ongoing) voice calls $N_i^{vg} + N_i^{vs} = n$, that is, when $l > K_v^g + K_v^s - n$, and the number of blocked calls is $l-(K_v^g+K_v^s-n)$. The stationary distribution of $P[N_i^{vg}+N_i^{vs} = n]$ is obtained through convolution from $\Pi^v(k,j)$, the joint distribution of (N_i^{vs}, N_i^{vg}), as

$$\lim_{i\to\infty} P[N_i^{vs} + N_i^{vg} = n] = \sum_{j=\max\{0,n-K_v^s\}}^{\min\{K_v^g,n\}} \Pi^v(n-j,j) \tag{3.12}$$

4. DATA TRAFFIC ANALYSIS

We first present the analysis of data process over the satellite subnet.

4.1. Data Traffic over the S-Subnet

At the beginning of frame i, all data nodes have information about the number of established voice calls in frame i. This information determines the number of slots available

for data traffic in frame i given by

$$L_i = L - 1 - N_i^{vs} \tag{4.1}$$

Let

Y_i^s = Number of data packet transmissions in frame i over S-subnet

C_i = Number of data packets involved in collision in frame i over S-subnet

S_i^s = Number of successful data packet transmissions in frame i over S-subnet

N_i^{ds} = Number of data packet retransmissions in frame i over S-subnet

A_i^d = Number of new data packet arrivals in frame i

Figure 5 depicts all components of data traffic over the satellite and ground subnets. The necessary notation for data traffic over the ground subnet is provided at the beginning of Section 4.2.

From the above definitions, we see that

$$Y_i^s = A_i^d + N_i^{ds} = S_i^s + C_i \tag{4.2}$$

It is assumed that all the data nodes that were involved in a collision event in frame i are able to estimate the number of packets that were involved in collision in frame i. This information becomes available at the end of frame $i+1$ because of the delay (= one frame = L slots) introduced by the satellite. Based on this information, a decision is made for each collided packet at the beginning of frame $i+2$. With probability γ_{i+2} it is transmitted over the S-subnet sometime during frame $i+3$ and with probability $1-\gamma_{i+2}$ it joins the backlog of data packets over the G-subnet. We will update the splitting factor γ_i, based on the fraction of data nodes that were involved in a collision event. This rule may be given by:

$$\gamma_{i+2} = \begin{cases} 1 & C_i \le \alpha M_d \\ \frac{\beta}{C_i} & C_i > \alpha M_d \end{cases} \tag{4.3}$$

where $\alpha > 0$* and $0 < \beta \le \alpha M_d$. * Clearly, increasing the values of either of these parameters results in an increase in the fraction of retransmissions that use the S-subnet.

* Note that, since each of the data users can transmit more than one packet per frame, it is possible to have $C_i > M_d$; thus $\alpha > 1$ may be appropriate in some cases.

The selection of the splitting probability $1 - \gamma_{i+2}$ above is not optimal. However, it is justified by the fact that, as the number of colliding data packets C_i increases, it is desirable to retransmit a larger portion of these packets over the ground subnet so that the the data traffic over the satellite subnet is somewhat relieved. One can obtain the optimal values of the parameters α and β involved in (4.3) through a numerical search for the optimization of performance criteria such as the data throughput and average data delay. One could also attempt to optimize over the form of the splitting factor γ_{i+2}, but this is a much more difficult and computationally demanding task which is beyond the scope of this report.

It can be shown that the evolution of the data process over the S-subnet can be described by the two-dimensional Markov process $\underline{D}_i^s \triangleq (C_i, N_i^{vs})$. The transition probability matrix of $\{\underline{D}_i^s\}$ may be obtained by observing that

$$\begin{aligned}
P\left[\underline{D}_{i+2}^s = (n,m) | \underline{D}_i^s = (k,j)\right] &= P\left[C_{i+2} = n | C_i = k, N_i^{vs} = j, N_{i+2}^{vs} = m\right] \\
&\quad \cdot P\left[N_{i+2}^{vs} = m | C_i = k, N_i^{vs} = j\right] \\
&= P\left[C_{i+2} = n | C_i = k, N_{i+2}^{vs} = m\right] \cdot P\left[N_{i+2}^{vs} = m | N_i^{vs} = j\right]
\end{aligned} \tag{4.4}$$

This decomposition follows from the fact that the voice process N_i^{vs} is independent of the data process and the number of data packets involved in a collision in frame $i + 2$ is independent of N_i^{vs} once C_i and N_{i+2}^{vs} are given.

The two-phase transition of N_i^{vs} can be obtained from the 2-step transition probability matrix and the stationary probability distribution of $\underline{N}_i^v$. The first term on the right hand side of (4.4) is given by

$$\begin{aligned}
P\left[C_{i+2} = n | C_i = k, N_{i+2}^{vs} = m\right] &= \sum_{l=n}^{M_d+k} P\left[C_{i+2} = n | C_i = k, N_{i+2}^{vs} = m, Y_{i+2}^s = l\right] \\
&\quad \cdot P\left[Y_{i+2}^s = l | C_i = k, N_{i+2}^{vs} = m\right] \\
&= \sum_{l=n}^{M_d+k} P\left[C_{i+2} = n | N_{i+2}^{vs} = m, Y_{i+2}^s = l\right] \cdot P\left[Y_{i+2}^s = l | C_i = k\right]
\end{aligned}$$

$$(4.5)$$

for $n \leq M_d + k$; the above quantity is 0 for $n > M_d + k$.

The second term in the sum of (4.5) is given by

$$P\left[Y_{i+2}^s = l | C_i = k\right] = \sum_{a=\max\{0,l-k\}}^{\min\{l,M_d\}} P\left[A_{i+2}^d = a | C_i = k\right] \cdot P\left[N_{i+2}^{ds} = l - a | C_i = k\right] \quad (4.6)$$

The conditional distributions of new arrivals and the retransmitted packets over the S-subnet are given by:

$$P\left[A_{i+2}^d = a | C_i = k\right] = b(M_d, a, P_d) \quad (4.7)$$

$$P\left[N_{i+2}^{ds} = l - a | C_i = k\right] = b(k, l - a, \gamma_{i+2}) \quad (4.8)$$

where γ_{i+2} is given by (4.3). Equation (4.7) is justified by the mechanism of data packet generation described in Section 2.2. Since the C_i data terminals with colliding packets can generate new packets in frame $i + 2$ (at the time they find out about the status of their transmitted packets in frame i) the number of data terminals that can generate new data packets is M_d independently of the value of C_i. On the other hand (4.8) is a direct consequence of the splitting scheme of (4.3).

The first term in the sum of (4.5) is equivalent to

$$\begin{aligned} P\left[C_{i+2} = n | N_{i+2}^{vs} = m, Y_{i+2}^s = l\right] &= P\left[S_{i+2}^s = l - n | N_{i+2}^{vs} = m, Y_{i+2}^s = l\right] \\ &= P\left[S_{i+2}^s = l - n | L_{i+2}^{vs} = L - 1 - m, Y_{i+2}^s = l\right] \quad (4.9) \end{aligned}$$

where the conditional distribution of S_i^s is given by:

$$P\left[S_i^s = k | Y_i^s = m, L_i = j\right] = \left(\frac{(-1)^k j! m!}{j^m k!}\right) \cdot \sum_{l=k}^{\min\{m,j\}} \frac{(-1)^l (j-l)^{m-l}}{(l-k)!(j-l)!(m-l)!} \quad (4.10)$$

Eq. (4.10) (taken from [9]) represents the probability that k cells are occupied with one ball (i.e. the successful data packet transmissions S_i^s) when m balls (i.e., the transmitted data packets Y_i^s) are thrown into j cells (i.e., the number of available slots L_i)

An important comment must be made here regarding (4.10) and the retransmission mechanism of the framed ALOHA protocol as applied to our mixed-media network model. Recall that, according to the model for the generation of new data packets described in Section 2.2, it is possible that data nodes generate new packets before several of their previously generated packets have been successfully transmitted. Thus, several data packets belonging to the same data node(s) may be part of N_i^{ds}, C_i, S_i^s, and Y_i^s of (4.2). Consequently, a number of data packets of the same data node may be competing for the L_i [given by (4.1)] available slots of the i-th frame. Under these conditions, any practical implementation of the framed ALOHA protocol will use some mechanism to avoid these suicidal collisions between packets belonging to the same data node by spreading the transmissions of these packets over several consecutive frames or by simply choosing different slots within (possibly) the same frame for its packets. This is less of a problem for larger values of M_d. However, the analysis of such a scheme is very complicated and beyond the scope of this report. Here, we obtain a lower bound on the probability distribution of successful data packets by pretending that all packets belong to different data nodes. Clearly, since any intelligent protocol would do better than this, (4.10) serves as a lower bound to the true probability and therefore all expressions about data throughput and delay based on (4.10) are pessimistic. For moderately small values of the data node activity parameter P_d, the probability of having more than one packet in the backlog N_i^{ds} belonging to the same data node is relatively small, so that the lower bound of (4.10) is relatively close to the actual value. For larger values of P_d this is not the case and further research is required to determine the tightness of the lower bound of (4.10) and the resulting loss in data throughput and delay performance.

Equations (4.4)-(4.10) provide us with the transition probability matrix of the two-dimensional Markov process $\underline{D}_i^s$. We may evaluate the stationary distribution of the process to obtain the throughput-delay performance of the data traffic over the satellite subnet.

The expected data throughput over the S-subnet is then given by

$$\eta_d^s = \sum_{N_i^{vs}=0}^{K_i^s} \sum_{N_i^{ds}=0}^{M_d} \sum_{S_i^s=0}^{L-1-N_i^{vs}} S_i^s P\left[S_i^s | N_i^{ds}, N_i^{vs}\right] P\left[N_i^{ds}, N_i^{vs}\right] \tag{4.11}$$

in packets per frame. In (4.11) the joint distribution $P\left[N_i^{ds}, N_i^{vs}\right]$ is obtained from

$$P\left[N_i^{ds}, N_i^{vs}\right] = P\left(N_i^{ds} \mid N_i^{vs}\right) P\left(N_i^{vs}\right) \tag{4.12}$$

where

$$\begin{aligned} P\left(N_i^{ds} | N_i^{vs}\right) &= \sum_{C_{i-2}=N_i^{ds}}^{\infty} P\left(N_i^{ds} \mid C_{i-2}, N_i^{vs}\right) P\left(C_{i-2} \mid N_i^{vs}\right) \\ &= \sum_{C_{i-2}=N_i^{ds}}^{\infty} P\left(N_i^{ds} \mid C_{i-2}\right) P\left(C_{i-2} \mid N_i^{vs}\right), \end{aligned} \tag{4.13}$$

$$P\left(N_i^{ds} = k \mid C_{i-2} = m\right) = b(m, k, \gamma_i), \tag{4.14}$$

and

$$\begin{aligned} P(C_{i-2} \mid N_i^{vs}) &= \frac{P(C_{i-2}, N_i^{vs})}{P(N_i^{vs})} \\ &= \frac{\sum_{C_i=0}^{\infty} \sum_{N_{i-2}^{vs}=0}^{K_v^s} P(C_i, N_i^{vs}, C_{i-2}, N_{i-2}^{vs})}{P(N_i^{vs})} \\ &= \frac{\sum_{C_i=0}^{\infty} \sum_{N_{i-2}^{vs}=0}^{K_v^s} P(C_i, N_i^{vs} \mid C_{i-2}, N_{i-2}^{vs}) \cdot P(C_{i-2}, N_{i-2}^{vs})}{P(N_i^{vs})} \end{aligned} \tag{4.15}$$

is obtained from the transition probability matrix and the stationary distribution of $\underline{D}_i^s = (C_i, N_i^{vs})$. Moreover, the conditional probability $P\left(S_i^s \mid N_i^{ds}, N_i^{vs}\right)$ is obtained from

$$\begin{aligned} P\left[S_i^s = n \mid N_i^{ds} = m, N_i^{vs} = k\right] = \sum_{l=\max\{m,n\}}^{M_d+m} & P\left[S_i^s = n \mid N_i^{ds} = m, N_i^{vs} = k, Y_i^s = l\right] \\ & \cdot P\left[Y_i^s = l \mid N_i^{ds} = m, N_i^{vs} = k\right] \end{aligned} \tag{4.16}$$

where

$$P\left[S_i^s = n \mid N_i^{ds} = m, N_i^{vs} = k, Y_i^s = l\right] = P\left[S_i^s = n \mid N_i^{vs} = k, Y_i^s = l\right]$$

is independent of the value of N_i^{ds} and is given by (4.9)-(4.10), and

$$P\left[Y_i^s = l \mid N_i^{ds} = m, N_i^{vs} = k\right] = P\left(A_i^d = l - m\right) = b(M_d, l - m, P_d) \tag{4.17}$$

Finally, the average data delay for a packet successfully transmitted over the S-subnet (measured in frames) is approximated as

$$\bar{D}^s \approx 1 + \frac{E\{\text{number of transmitted data packets per frame}\}}{E\{\text{number of successfully transmitted packets per frame}\}}$$

$$= 1 + \frac{1}{\eta_d^s} \cdot \lim_{i\to\infty} E\{Y_i^s\}$$

$$= 1 + \frac{1}{\eta_d^s} \cdot \lim_{i\to\infty} \sum_{N_i^{vs}} \sum_{N_i^{ds}} \sum_{Y_i^s} Y_i^s P\left[Y_i^s | N_i^{ds}, N_i^{vs}\right] P\left[N_i^{ds}, N_i^{vs}\right]$$

$$= 1 + \frac{1}{\eta_d^s} \cdot \lim_{i\to\infty} \sum_{k=0}^{K_v^s} \sum_{m=0}^{\infty} \sum_{l=m}^{M_d+m} l P\left[Y_i^s = l | N_i^{ds} = m, N_i^{vs} = k\right] P\left[N_i^{ds} = m, N_i^{vs} = k\right]$$

(4.18)

where $P\left[Y_i^s = l | N_i^{ds} = m, N_i^{vs} = k\right]$ is obtained from (4.17) and $P\left[N_i^{ds} = m, N_i^{vs} = k\right]$ is obtained from (4.12). The term 1 in the right hand side of (4.18) accounts for the fixed delay of one frame (= L slots) introduced by the satellite to all data packets transmitted through it.

Notice that the expression in (4.18) is only an approximation to the actual data delay over the S-subnet. The method followed for this derivation parallels the approximate delay analysis provided in [14, Section 4.2, pp. 219-221] for slotted multiple-access systems with retransmissions. Little's law is not exactly used in (4.18), but rather a similar intuitive argument produces the first equation in (4.18); refer to [14] for a detailed description of this approximation. Our situation is different from that of [14] in that we are interested in the overall system delay (waiting time + service time) whereas [14] deals with waiting time (queuing delay) only; of course the protocol of [14] is also different.

4.2. Data Traffic over the *G*-Subnet

Some of the data nodes that were involved in a collision even over the S-subnet choose to retransmit the collided packets over the ground subnet.

Let

R_i^{dg} = Number of data packets in backlog over G-subnet at the beginning of frame i

N_i^{dg} = Number of data packets from S-subnet that join backlog of data over G-subnet at the beginning of frame i

S_i^g = Number of successfully transmitted data packets over G-subnet in frame i

T_i = Number of packets that actually transmit over G-subnet in frame i

Figure 5 depicts all components of data traffic over the ground subnet.

It is easy to see that

$$R_{i+1}^{dg} = R_i^{dg} + N_i^{dg} - S_i^g \tag{4.19}$$

The protocol observed by the data terminals over the G-subnet is as follows: At the beginning of frame i, the terminals with packets to transmit in the G-subnet have information about the number of established voice calls and the data backlog over the G-subnet. The total number of these data terminals is $R_i^{dg} + N_i^{dg}$; each of these data nodes transmits its packet in frame i with probability f_i given by

$$f_i = \begin{cases} \min\left\{1, \frac{L'(N_i^{vg})}{N_i^{dg}+R_i^{dg}}\right\}; & \text{if } L'(N_i^{vg}) > 0 \\ 0; & \text{if } L'(N_i^{vg}) \leq 0 \end{cases} \tag{4.20}$$

where

$$L'(N_i^{vg}) = K_d^g - N_i^{vg} \tag{4.21}$$

denotes the unused multiple-access capability of the CDMA channel of the G-subnet in frame i; this channel capacity is available for use by the T_i (out of the $R_i^{dg} + N_i^{dg}$) data users that actually transmit.

The selection of f_i in (4.20) as the retransmission probability is well justified by the fact that it makes the average number of transmitting users equal to the available CDMA capacity, that is,

$$E\{T_i\} = f_i(R_i^{dg} + N_i^{dg}) = L'(N_i^{dg}) = K_d^g - N_i^{vg}$$

This type of retransmission probabilities were first introduced in [10] and [11] in the context of ALOHA protocols with retransmission control. We have also used them in [6] for the retransmission control of multi-media (voice/data) CDMA networks. Actually, several

variants of (4.20) were used together with different feedback information for the schemes termed A, B, and C whose performance was analyzed in [6].

Let $\underline{D}_i^g = (R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1})$. Then

$$P[\underline{D}_{i+1}^g = (m_1, m_2, m_3, m_4) | \underline{D}_i^g = (k_1, k_2, k_3, k_4)]$$

$$= P[R_{i+1}^{dg} = m_1 \mid N_{i+1}^{dg} = m_2, N_{i+1}^{vg} = m_3, C_i = m_4, R_i^{dg} = k_1, \ N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4]$$

$$\cdot P[N_{i+1}^{dg} = m_2 \mid N_{i+1}^{vg} = m_3, C_i = m_4, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4]$$

$$\cdot P[N_{i+1}^{vg} = m_3 \mid C_i = m_4, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4]$$

$$\cdot P[C_i = m_4 \mid R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \tag{4.22}$$

For the third of the four terms in the sum on the right hand side of (4.22), since voice evolves independently of data, we have:

$$P[N_{i+1}^{vg} = m_3 \mid C_i = m_4, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] = P[N_{i+1}^{vg} = m_3 \mid N_i^{vg} = k_3], \tag{4.23a}$$

which is obtained from the joint conditional distribution $P(N_{i+1}^{vs}, N_{i+1}^{vg} | N_i^{vs}, N_i^{vg})$ and the stationary distributions $P(N_i^{vs}, N_i^{vg})$ and $P(N_i^{vg})$ evaluated in Section 3 as

$$P(N_{i+1}^{vg} | N_i^{vg}) = \frac{P(N_{i+1}^{vg}, N_i^{vg})}{P(N_i^{vg})}$$

$$= \frac{\sum_{N_{i+1}^{vs}=0}^{K_v^s} \sum_{N_i^{vs}=0}^{K_v^s} P(N_{i+1}^{vs}, N_{i+1}^{vg}, N_i^{vs}, N_i^{vg})}{P(N_i^{vg})}$$

$$= \frac{\sum_{N_{i+1}^{vs}=0}^{K_v^s} \sum_{N_i^{vs}=0}^{K_v^s} P(N_{i+1}^{vs}, N_{i+1}^{vg} | N_i^{vs}, N_i^{vg}) \cdot P(N_i^{vs}, N_i^{vg})}{\sum_{N_i^{vs}=0}^{K_v^s} P(N_i^{vs}, N_i^{vg})} \tag{4.23b}$$

Moreover, for the second term on the right hand side of (4.22) we have

$$P[N_{i+1}^{dg} = m_2 \mid N_{i+1}^{vg} = m_3, C_i = m_4, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4]$$

$$= P[N_{i+1}^{dg} = m_2 \mid C_{i-1} = k_4]$$

$$= b(k_4, m_2, 1 - \gamma_{i+1}). \tag{4.24}$$

In deriving (4.24) we used the fact that that N_{i+1}^{dg}, the number of data packets from the satellite subnet joining the backlog of the ground subnet during frame $i+1$, depends only

on C_{i-1}, the number of colliding data packet over the satellite subnet during frame $i-1$; refer to (4.3) for the expression determining the splitting probability γ_{i+1}.

The fourth term on the right-hand side of (4.22) is given by

$$
\begin{aligned}
&P[C_i = m_4 \mid R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \\
&= \sum_{N_i^{vs}=0}^{K_v^s} \sum_{C_{i-2}=0}^{\infty} P[C_i = m_4 \mid N_i^{vs}, C_{i-2}, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \\
&\qquad \cdot P[C_{i-2} \mid N_i^{vs}, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \\
&\qquad \cdot P[N_i^{vs} \mid R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \\
&= \sum_{N_i^{vs}=0}^{K_v^s} \sum_{C_{i-2}=k_2}^{\infty} P[C_i = m_4 \mid N_i^{vs}, C_{i-2}] \cdot P[C_{i-2} \mid N_i^{dg} = k_2] \cdot P[N_i^{vs} \mid N_i^{vg} = k_3]
\end{aligned}
\tag{4.25a}
$$

where the first term on the right hand side of (4.25a) can be obtained from (4.5), the second term can be obtained from (4.24) and is given below as

$$
\begin{aligned}
&P(C_{i-2} \mid N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1}) \\
&= \frac{P(C_{i-2}, N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1})}{P(N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1})} \\
&= \frac{P(N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1} | C_{i-2}) \cdot P(C_{i-2})}{P(N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1})} \\
&= \frac{P(N_i^{dg} | C_{i-2}, N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1}) \cdot P(N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1} | C_{i-2}) \cdot P(C_{i-2})}{P(N_i^{vs}, R_i^{dg}, N_i^{dg}, N_i^{vg}, C_{i-1})} \\
&= \frac{P(N_i^{dg} | C_{i-2}) \cdot P(N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1} | C_{i-2}) \cdot P(C_{i-2})}{P(N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1} | N_i^{dg}) \cdot P(N_i^{dg})} \\
&= \frac{P(N_i^{dg} | C_{i-2}) \cdot P(N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1}) \cdot P(C_{i-2})}{P(N_i^{vs}, R_i^{dg}, N_i^{vg}, C_{i-1}) \cdot P(N_i^{dg})} \\
&= \frac{P(N_i^{dg} \mid C_{i-2}) \cdot P(C_{i-2})}{P(N_i^{dg})}
\end{aligned}
$$

$$= P(C_{i-2} \mid N_i^{dg})$$

$$= \frac{P(N_i^{dg} \mid C_{i-2}) \cdot P(C_{i-2})}{\sum_{C_{i-2}=0}^{\infty} P(N_i^{dg} | C_{i-2}) \cdot P(C_{i-2})}$$

$$= \frac{b(C_{i-2}, N_i^{dg}, 1-\gamma_i) \cdot P(C_{i-2})}{\sum_{C_{i-2}=N_i^{dg}}^{\infty} b(C_{i-2}, N_i^{dg}, 1-\gamma_i) \cdot P(C_{i-2})} \tag{4.25b}$$

where $P(C_{i-2})$ is obtained from the Markov chain (C_i, N_i^{vs}) of Section 4.1, and the last term on the right-hand side of (4.25a) is given as

$$P(N_i^{vs} \mid N_i^{vg}) = \frac{P(N_i^{vs}, N_i^{vg})}{P(N_i^{vg})} = \frac{P(N_i^{vs}, N_i^{vg})}{\sum_{N_i^{vs}=0}^{K_v^s} P(N_i^{vs}, N_i^{vg})} \tag{4.25c}$$

Finally, for the first term in the sum on the right hand side of (4.22), we have two options for the model of correct reception of data packets: the threshold model and the graceful degradation model. If the **threshold model** of Section 2.1 is used for the reception of data packets, according to which data packets are successful (i.e., correctly received) if the total number of simultaneously transmitted packets (data and voice) over the G-subnet is smaller than K_d^g, we have

$$P[R_{i+1}^{dg} = m_1 \mid N_{i+1}^{dg} = m_2, N_{i+1}^{vg} = m_3, C_i = m_4, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4]$$

$$= P[R_{i+1}^{dg} = m_1 | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3]$$

$$= \sum_l P[R_{i+1}^{dg} = m_1 | T_i = l, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3] \cdot P[T_i = l | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3]$$

$$= \begin{cases} \left[\sum_{l=L'(k_3)+1}^{k_1+k_2} b(k_1+k_2, l, f_i)\right] + b(k_1+k_2, 0, f_i) & \text{if } m_1 = k_1+k_2,\ 0 < L'(k_3) < k_1+k_2 \\ 1 & \text{if } m_1 = k_1+k_2,\ L'(k_3) \le 0 \\ b(k_1+k_2, k_1+k_2-m_1, f_i) & \text{if } \max\{0, k_1+k_2-L'(k_3)\} \le m_1 < k_1+k_2 \\ 0 & \text{elsewhere} \end{cases} \tag{4.26a}$$

where $L'(k_3) = K_d^g - k_3$ as defined by (4.21).

On the other hand, we may use the **graceful degradation model** for CDMA multiple-reception of data packets (see [10]), according to which there is a non-zero probability $P(n|k)$ of correct reception for any arbitrary number n of packets and k is the total number of simultaneously transmitted packets (data and voice) over the G-subnet. In this case we have

$$
\begin{aligned}
&P[R_{i+1}^{dg} = m_1 \mid N_{i+1}^{dg} = m_2, N_{i+1}^{vg} = m_3, C_i = m_4, R_i^{dg} = k_1,\ N_i^{dg} = k_2, N_i^{vg} = k_3, C_{i-1} = k_4] \\
&= P[R_{i+1}^{dg} = m_1 | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3] \\
&= \sum_l P[R_{i+1}^{dg} = m_1 | T_i = l, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3] \cdot P[T_i = l | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3] \\
&= \sum_{l=0}^{k_1+k_2} P(k_1+k_2-m_1 \text{ successful data packets} | l+k_3 \text{ total transmitted packets})\, b(k_1+k_2, l, f_i)
\end{aligned}
\tag{4.26b}
$$

where

$$
\begin{aligned}
&P(n \text{ successful data packets} \mid l + k_3 \text{ total transmitted packets}) \\
&\qquad \approx \begin{cases} \binom{l}{n} [1 - P_E(l + k_3)]^n \, [P_E(l + k_3)]^{l-n} & \text{if } 0 \le n \le l \\ 0 & \text{elsewhere} \end{cases}
\end{aligned}
\tag{4.27}
$$

and the probability of packet error for a frequency-hopped CDMA system using Reed-Solomon (N, K) coding and erasures-decoding is given for k simultaneous packet transmissions by

$$
P_E(k) = \sum_{i=N-K+1}^{N} \binom{N}{i} [P_h(k)]^i \, [1 - P_h(k)]^{N-i} \tag{4.28}
$$

where the probability of a hit from other users is given by [11]

$$
P_h(k) = \begin{cases} 1 - \left(1 - \frac{1}{q}\right)^{k-1} & \text{for hop synchronous FH systems} \\ 1 - \left(1 - \frac{2}{q}\right)^{k-1} & \text{for hop asynchronous FH systems} \end{cases} \tag{4.29}
$$

The approximation in (4.27), termed the independence of receiver operation assumption (IROA), has been shown in [10] to have satisfactory accuracy for most situations of practical interest. The exact expression for the quantity in the left hand side of (4.27) has computation complexity that is typically exponential in n (e.g., $2^n - 1$ sums are involved in the exact expression for $P(n|l)$ for FH/SS systems with RS coding and erasures-decoding); refer to [10] for a detailed discussion of these issues and numerical examples establishing the accuracy of the approximation of (4.27). Finally, recall that our system is packet (= frame) synchronous in the ground subnet and slot synchronous in the satellite subnet; however, it is not required that the FH/SS system employed in the ground subnet is synchronous at the hop (or dwell time) level; this is the reason for providing two expressions for the probability of a hit in (4.29).

Finally, notice that the number of data packets in backlog at the beginning of frame i over the G-subnet, R_i^{dg}, has a countably infinite space; since it is possible (although with very small probability) that several of the data packets that get transmitted with probability f_i are not actually transmitted or they collide repeatedly and thus according to our packet arrival model data packets may accumulate indefinitely in the buffers of the data nodes. Therefore, in the numerical evaluation we must truncate the transition probability matrix of (4.22) and (4.26) to some moderately large value of R_i^{dg}.

The expected data throughput over the G-subnet can be computed as

$$\eta_d^g = \sum_{N_i^{vg}=0}^{K_d} \sum_{R_i^{dg}=0}^{\infty} \sum_{N_i^{dg}=0}^{\infty} \sum_{S_i^g=0}^{L'(N_i^{vg})} S_i^g P(S_i^g \mid R_i^{dg}, N_i^{dg}, N_i^{vg}) P(R_i^{dg}, N_i^{dg}, N_i^{vg}) \tag{4.30}$$

in packets/frame and for the **threshold model** of CDMA other-user interference

$$\begin{aligned} &P(S_i^g = n | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \\ &\quad = \sum_l P(S_i^g = n | T_i = l, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \cdot P(T_i = l | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \end{aligned}$$

$$
= \begin{cases} \left[\sum_{l=L'(k_3)+1}^{k_1+k_2} b(k_1+k_2,l,f_i)\right] + b(k_1+k_2,0,f_i) & \text{if } n=0,\ \ 0 < L'(k_3) < k_1+k_2 \\ 1 & \text{if } n=0,\ \ L'(k_3) \le 0 \\ b(k_1+k_2,n,f_i) & \text{if } 0 < n \le \min\{L'(k_3), k_1+k_2\} \\ 0 & \text{elsewhere} \end{cases} \tag{4.31a}
$$

where $L'(k_3)$ is defined as in (4.21). If the **graceful degradation model** for CDMA other-user interference is employed we should use

$$
\begin{aligned}
&P(S_i^g = n | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \\
&\quad = \sum_l P(S_i^g = n | T_i = l, R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \\
&\qquad\qquad \cdot P(T_i = l | R_i^{dg} = k_1, N_i^{dg} = k_2, N_i^{vg} = k_3) \\
&\quad = \sum_{l=0}^{k_1+k_2} P(n \text{ successful data packets} \mid l + k_3 \text{ total transmitted packets})\, b(k_1+k_2, l, f_i)
\end{aligned} \tag{4.31b}
$$

where $P(n \text{ successful data packets} \mid l + k_3 \text{ total transmitted packets})$ is given by (4.27)-(4.29). Finally, on the right hand side of (4.30) $P(R_i^{dg} = j, N_i^{dg} = m, N_i^{vg} = n)$ denotes the stationary distribution of the Markov chain $\underline{D}^g$ whose one-step transition probabilities were evaluated above.

Finally, the expected data delay of a packet successfully transmitted over the G-subnet (measured in frames) is given by

$$
\bar{D}^g = D^{gs} + D^{gg} \tag{4.32}
$$

where $\bar{D}^{gs}$ and $\bar{D}^{gg}$ denote the components of the delay of a typical packet transmitted over the G-subnet, which are due to protocol operation over the S-subnet and G-subnet, respectively, and are given by

$$\bar{D}^{gs} \approx 1+\frac{E\{\text{number of colliding data packets per frame in S} - \text{subnet}\}}{E\{\text{number of data packets per frame joining the G} - \text{subnet backlog}\}}$$

$$= 1+\lim_{i\to\infty}\frac{E\{C_i\}}{E\{N_i^{dg}\}}$$

$$= 1+\lim_{i\to\infty}\frac{\sum_{N_i^{vs}=0}^{K_v^s}\sum_{N_i^{ds}=0}^{\infty}\sum_{C_i=N_i^{ds}}^{\infty} C_i P\left[C_i|N_i^{vs},N_i^{ds}\right] P\left[N_i^{vs},N_i^{ds}\right]}{\sum_{R_i^{dg}=0}^{\infty}\sum_{N_i^{dg}=0}^{\infty}\sum_{N_i^{vg}=0}^{K_v^g} N_i^{dg} P[R_i^{dg},N_i^{dg},N_i^{vg}]}$$

$$= 1+\lim_{i\to\infty}\frac{\sum_{N_i^{vs}=0}^{K_v^s}\sum_{N_i^{ds}=0}^{\infty}\sum_{C_i=0}^{\infty}\sum_{Y_i^s=0}^{\infty} C_i P\left[N_i^{vs},N_i^{ds}\right] P\left[C_i|N_i^{vs},N_i^{ds},Y_i^s\right] P\left[Y_i^s|N_i^{vs},N_i^{ds}\right]}{\sum_{R_i^{dg}=0}^{\infty}\sum_{N_i^{dg}=0}^{\infty}\sum_{N_i^{vg}=0}^{K_v^g} N_i^{dg} P[R_i^{dg},N_i^{dg},N_i^{vg}]}$$

$$= 1+\lim_{i\to\infty}\frac{\sum_{N_i^{vs}=0}^{K_v^s}\sum_{N_i^{ds}=0}^{\infty}\sum_{C_i=0}^{\infty}\sum_{Y_i^s=0}^{\infty} C_i P\left[N_i^{vs},N_i^{ds}\right] P\left[C_i|N_i^{vs},Y_i^s\right] P\left[Y_i^s|N_i^{ds}\right]}{\sum_{R_i^{dg}=0}^{\infty}\sum_{N_i^{dg}=0}^{\infty}\sum_{N_i^{vg}=0}^{K_v^g} N_i^{dg} P[R_i^{dg},N_i^{dg},N_i^{vg}]}$$

$$= 1+\lim_{i\to\infty}\frac{\sum_{m=0}^{K_v^s}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\sum_{l=\max\{k,n\}}^{M_d+k} nP\left[N_i^{vs}=m,N_i^{ds}=k\right]P[C_i=n|N_i^{vs}=m,Y_i^s=l]b(M_d,l-k,P_d)}{\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{K_v^g} n_2 P[R_i^{dg}=n_1,N_i^{dg}=n_2,N_i^{vg}=n_3]}$$

$$(4.33)$$

and

$$\bar{D}^{gg} \approx 1+\frac{E\{\text{number of backlogged data packets per frame in G} - \text{subnet}\}}{E\{\text{number of successful data packets per frame in G} - \text{subnet}\}}$$

$$= 1+\frac{1}{\eta_d^g}\lim_{i\to\infty}\left[E\{R_i^g\}+E\{N_i^g\}\right]$$

$$= 1+\frac{1}{\eta_d^g}\lim_{i\to\infty}\sum_{R_i^{dg}=0}^{\infty}\sum_{N_i^{dg}=0}^{\infty}[R_i^{dg}+N_i^{dg}]P(R_i^{dg},N_i^{dg}) \qquad (4.34).$$

In (4.33), $P\left[N_i^{vs}=m,N_i^{ds}=k\right]$ and $P\left[C_i=n|N_i^{vs}=m,Y_i^s=l\right]$ are obtained from (4.12) and (4.9)-(4.10), respectively. Similarly, in (4.34), $P(R_i^{dg},N_i^{dg})$ can be evaluated from $P(R_i^{dg},N_i^{dg},N_i^{vg})$, the stationary distribution of the Markov chain $\underline{D}_i^g=(R_i^{dg},N_i^{dg},N_i^{vg})$. Moreover, the term 1 on the right hand side of (4.33) accounts for the fixed satellite delay of one frame that any data packet transmitted over the satellite is subject to. The second term in (4.33) represents the component of the delay due to collisions and retransmissions

over the S-subnet before a data packet gets retransmitted over the G-subnet. The term 1 in the right hand side of (4.34) accounts for the fact that data packets from the S-subnet retransmitted through the G-subnet are delayed by an additional one frame period [because of the delay in the satellite channel feedback providing information about packet collisions]. Finally, the second term in the right hand side of (4.34) represents the component of the delay due to collisions and retransmissions over the G-subnet.

Finally, notice that the expressions in (4.33) and (4.34) are only approximations to the actual delay terms. The method followed for their derivations is similar to that used for equation (4.18) (data delay over the S-subnet) and parallels the approximate delay analysis provided in [14, Section 4.2, pp. 219-221] for slotted multiple-access systems with retransmissions.

5. NUMERICAL RESULTS

In all figures with numerical results presented in this section, the parameters of voice and data traffic, as well as the capacities of the ground and satellite channels, are depicted on each figure.

The offered voice traffic G_v is defined as

$$G_v = \lim_{i \to \infty} \sum_{n=0}^{K_v^g + K_v^s} [(M_v - n)P_v + n] P\left[N_i^{vg} + N_i^{vs} = n\right]$$

$$= M_v P_v + (1 - P_v) \sum_{n=0}^{K_v^g + K_v^s} n \lim_{i \to \infty} P\left[N_i^{vg} + N_i^{vs} = n\right] \qquad (5.1)$$

where $\lim_{i \to \infty} P\left[N_i^{vg} + N_i^{vs} = n\right]$ is given by (3.12).

The basic set of system and traffic parameters that is used in all figures presented below is: $M_v = 20$ and $M_d = 15$ for the total population of voice and data users, respectively: $K_v^g = 7$ and $K_d^g = 5$ for the multiple-access capability of the CDMA ground network [refer to (2.1)-(2.2)] for voice and data users, respectively; $p = 0.05$ for the parameter of the geometric distribution governing the length of the voice calls [refer to (2.3)]; and $L = 13$ for the frame length, which is equal to the constant delay introduced by the satellite.

Voice throughput, defined as the expected total number of voice calls in progress per frame (summed over the G- and S-subnets), is shown in Figure 6 as a function of the offered

voice traffic G_v defined by (5.1). All voice calls are accepted when the voice traffic load is light. For a heavy voice traffic load, the throughput reaches saturation as the number of new voice calls increases, since most of the new calls are blocked. The traffic level at which saturation is reached depends on K_v^s, the number of slots per frame in the satellite channel that are available for voice traffic. Clearly, the availability of time slots in the satellite channel increases the voice throughput of the overall system.

Figure 7 illustrates the blocking probability of voice calls in the satellite and ground networks, as a function of the offered voice traffic G_v. As we allocate a larger part of the satellite channel capacity to voice traffic (by increasing K_v^s), blocking performance improves. However, an increase in K_v^s results in a decrease in the number of slots available for data traffic in the satellite subnet, and thus in the degradation of the data performance measures.

Average data throughput over the satellite subnet (S-subnet) is shown in Figure 8 as a function of the duty factor P_d of the data nodes. In Figure 8, when voice traffic is light, as is for the case ($G_v = 3.6$, $P_v = 0.01$, $K_v^s = 5$), there are more channels available for data traffic, fewer collisions occur and throughput over the S-subnet initially increases and then starts decreasing slowly as P_d increases beyond the value 0.35. due to increasing data packet collisions. On the other hand, when voice traffic is heavier, as is for the case ($G_v = 11.7$, $P_v = 0.1$, $K_v^s = 5$), data throughput, after reaching its (significantly lower) maximum value, decreases due to an increase in the number of colliding packets. Finally, for the case of substantially heavier voice traffic ($G_v = 18$, $P_v = 0.5$, $K_v^s = 10$), there are so few channels available for data traffic that the data throughput decreases slowly over the entire range of values of P_d. Changing the routing parameters α and β [refer to equation (4.3)] does not significantly affect the data throughput for either light or heavy voice traffic. It appears that for light voice traffic ($G_v = 3.6$) there is some loss in the data throughput if values of α and β that divert more data traffic to the G-subnet (e.g., $\alpha = 1.0$ and $\beta = 3.0$ versus $\alpha = 0.6$ and $\beta = 8.0$) are used; this observation is valid for the entire range of values of P_d. As the offered voice traffic load increases to $G_v = 11.7$ the pair $(\alpha, \beta) = (0.6, 8.0)$ results in slightly larger data throughput than the pair $(\alpha, \beta) = (1.0, 3.0)$ for values of P_d smaller than 0.35, whereas the opposite is true for larger values of P_d. Finally, for

heavy voice traffic ($G_v = 18.0$), the pair $(\alpha, \beta) = (1.0, 3.0)$ results in slightly larger data throughput than the pair $(\alpha, \beta) = (0.6, 8.0)$ for all values of P_d; that is, it is advantageous to divert more data traffic to the G-subnet.

The parameters α and β of (4.3) determine the value of the splitting factor $1 - \gamma_{i+2}$, that is, the fraction of the data traffic of the S-subnet involved in packet collisions (i.e., C_i) which is diverted to the G-subnet. From these two parameters, α determines the threshold (which is equal to a fraction of the overall number of data users) which should be surpassed by C_i for data traffic to be diverted to the G-subnet, whereas β determines the actual value of the splitting factor as β / C_i. The effect of the splitting factor becomes more pronounced as α and β increase because in this case both the threshold and the value of the splitting factor increase. In most situations, larger values of α cause larger increases in $1 - \gamma_{i+2}$ than larger values of β.

Figure 9 displays the average data delay over the S-subnet as a function of the duty factor of the data nodes for the same system and traffic parameters as in Figure 8. Similar trends as in Figure 8 are observed. For light voice traffic the data delay is small and is almost insensitive to the changes in the values of α and β. However, for heavier voice traffic the data delay increases and becomes more dependent on the values of α and β; actually, as more data traffic is diverted through the G-subnet $[(\alpha, \beta) = (1.0, 3.0)$ versus $(\alpha, \beta) = (0.6, 8.0)]$ the data delay over the S-subnet improves significantly.

In Figures 10 and 11, the data throughput and data delay over the ground subnet (G-subnet) are illustrated as functions of the duty factor of the data nodes P_d for the threshold model described in Section 2.1. The parameters of the threshold model for the ground subnet are $K_v^g = 7$ and $K_d^g = 5$. These numbers correspond to the multiple-access capability of a specific FH/SS system (with parameters to be provided below) for desirable packet error probabilities of approximately $P_E^v = 10^{-3}$ (for voice) and $P_E^d = 10^{-5}$ (for data) in (2.1) and (2.2), respectively. The aforementioned FH/SS system employs 32-ary ($M = 32$) FSK modulation with noncoherent demodulation, 40 hopping frequencies [i.e., $q = 40$ in (4.29)], and an RS (32,16) code [i.e., $N = 32, K = 16$ in (4.28)] with erasures-decoding; perfect side information is assumed available from the channel, that is, all erasures are only due to multiple-access interference and can be detected without

errors. The other system and traffic parameters are the same as the ones used in Figure 8.

In Figure 10 in which the data throughput over the G-subnet is plotted versus P_d, when voice traffic is light, the number of channels available for data increases so that throughput increases as P_d increases to 0.3 and then decreases slightly. On the other hand, when voice traffic is heavier, data traffic can use only the channels that are not used for voice calls in the G-subnet. Data throughput thus decreases faster, because the number of colliding data packets increases whenever voice traffic (G_v) and data traffic (P_d) increase. Comparison of the throughput for two different pairs of the parameters (α, β) shows that the pair $(\alpha, \beta) = (0.6, 8.0)$ results in better performance than the pair $(1.0, 3.0)$ suggesting that an overall smaller splitting factor $(1 - \gamma_i)$ is preferable because fewer data packet collisions occur in the G-subnet (due to fewer packets being diverted from the S-subnet) and more data traffic goes through successfully.

In Figure 11, the average data delay over the G-subnet (which includes the delay experienced while unsuccessfully attempting transmission over the satellite subnet) is displayed as a function of the duty factor of the data nodes. The same threshold model as for Figure 10 is used. In general, data delay increases as voice traffic becomes heavier over the G-subnet. The splitting factor has a significant effect only when the voice traffic is heavier, in which case larger splitting factors cause an increase in data delay.

In Figures 12 and 13, the average data throughput and data delay over the G-subnet are shown as functions of the duty factor of the data nodes for the graceful degradation model. This model is described in detail in Section 4.2 [see equations (4.27)-(4.29)]. The same FH/SS system model parameters used for the threshold model (refer to the discussion about Figures 10 and 11 above) are used here as well. However, to reduce the computational burden in the graceful degradation model it has been necessary to truncate $P(n \mid l + k_3)$ of (4.27). In particular, in the process of generating numerical results we evaluate $P(n \mid l + k_3)$ from (4.27)-(4.29) only for the range $0 \leq l \leq 10$ and set $P(n \mid l + k_3) = 0$ for $l > 10$ for all $0 \leq k_3 \leq K_v^g$. For $k_3 = K_v^g = 7$, this corresponds to setting the packet error probability [of (4.28)] $P_E(l + k_3) = 1$ for all values of $P_E(l + k_3)$ larger than 0.7. For $k_3 = 0$, we set $P_E(l + k_3) = 1$ for all values of $P_E(l + k_3)$ larger than 0.08. For each intermediate value of k_3 $(0 < k_3 < K_v^g)$ the truncation (that is forcing $P_E(l + k_3) = 1$) takes place at some

different value of $P_E(l + k_3)$ between 0.08 and 0.7. Since large values (close to 1) of the packet error probabilities $P_E(l + k_3)$ result in extremely small probabilities $P(n \mid l + k_3)$, truncating the graceful degradation model in this way does not significantly reduce its accuracy for high voice traffic, because in this case k_3 takes larger values (closer to K_v^g) with high probability. However, for low voice traffic, in which case k_3 takes smaller values with high probability, the accuracy of the graceful degradation model decreases and the results of any comparison with the threshold model are expected to be pessimistic.

Since the graceful degradation model described above is computationally very demanding, in Figures 12 and 13 we present results for only two values of the average offered voice traffic per frame G_v (light: $G_v = 3.6$, $P_v = 0.01$, $K_v^s = 5$ and heavy: $G_v = 11.7$, $P_v = 0.1$, $K_v^s = 5$); the values of α and β are 1.0 and 3.0, respectively. It is observed that the behavior of the data throughput and delay is basically similar to that observed in Figures 10 and 11 for the threshold model for the case of light voice traffic. However, for the case of heavy voice traffic, the graceful degradation model shows more drastic decrease in the throughput and increase in the delay as P_d increases beyond the value .4. Comparison of the two curves in Figure 12 establishes that for small values of P_d data throughput is greater for heavy voice traffic than for light voice traffic. This is justified by the fact that more data packets use the G-subnet in the former case than in the latter case due to the increased number of data packets diverted from the S-subnet to the G-subnet.

In Figures 14 and 15, the average data throughput and data delay over the G-subnet are plotted versus the duty factor of the data nodes for both the threshold and the graceful degradation models. These comparisons are carried out for low (3.6) and high (11.7) average values of the offered voice traffic per frame G_v. The results of the comparison of data throughput in Figure 14 are different for the low and high G_v values. For low G_v, using the threshold model results in better performance than using the graceful degradation model, for all values of P_d, the duty factor of the data nodes. This can be explained by the fact that the truncation in the graceful degradation model discussed earlier is more harmful to it for low voice traffic. If such truncation were not necessary, the graceful degradation model would perform better than the threshold model for all cases of interest. By contrast, for high G_v, using the graceful degradation model results in better performance than using

the threshold model for P_d larger than 0.25, whereas the use of two models gives very similar results for P_d smaller than 0.25. Moreover, in Figure 15, the data delay of the threshold model appears to be significantly larger than the delay of the graceful degradation model for both low and high G_v; this ordering of data delays is valid for all values of the data duty factor P_d except for values smaller than .15, in which case the delay under the two models is nearly the same for light voice traffic.

In Figures 16 and 17, the average data delay over the S-subnet ($\bar{D}^s$), the average delay over the G-subnet ($\bar{D}^g$), and the the total average delay of a typical data packet in the system (transmitted over either of the two subnets) $\bar{D}$, defined as

$$\bar{D} = \frac{\eta_d^s}{\eta_d^s + \eta_d^g} \cdot \bar{D}^s + \frac{\eta_d^g}{\eta_d^s + \eta_d^g} \cdot \bar{D}^g. \tag{5.2}$$

are shown as functions of the duty factor of the data nodes P_d for the cases of light voice traffic ($G_v = 3.6, P_v = 0.01$) and heavy voice traffic ($G_v = 11.7, P_v = 0.1$), respectively. The various delays are displayed for both the threshold and graceful degradation models. Clearly, in regard to comparisons, the total average delay $\bar{D}$ obtained by using the two different models (threshold and graceful degradation) of correct data packet reception follows the same ordering as the data delays over the G-subnet discussed in Figure 15; however, the difference in the values of the total delay under the two models is less pronounced compared to that of the delays over the G-subnet.

6. CONCLUSIONS

In this report, we have studied the performance of certain multiple-access strategies for voice/data integration in heterogeneous mixed-media packet radio networks. Our models, protocols, and performance analysis are applicable to networks with terrestrial and satellite links. These multiple-access schemes take advantage of different characteristics of the ground and satellite radio channels to better serve the requirements of voice and data traffic. In the ground subnetwork, CDMA with movable boundary in the code domain is used to serve the voice traffic and a fraction of the retransmitted data traffic, while framed ALOHA with movable boundary is used on the satellite subnetwork for data and reservations in each frame are used for voice. In our scheme, data nodes attempt to

transmit over the satellite network first and the amount of data traffic transferred to the ground network depends on the number of colliding data packets over the satellite network.

An accurate Markovian model of the evolution of the voice and data traffic over the ground and satellite subnetworks has been developed involving two- and four-dimensional Markov chains. Moreover, two models for the successful reception of data packets are used: a threshold model and a graceful degradation model. On the basis of the above models, the steady-state performance of the two subnetworks is evaluated in terms of the blocking probability of voice calls, the average voice throughput (number of calls in progress per frame), the average data throughput, and the average data delay.

The performance evaluation of the multiple-access schemes considered in this report has established the interplay between different traffic types at different offered traffic levels. It has also established the sensitivity of data throughput and delay to the proper selection of the parameters determining the amount of retransmitted data traffic routed through the ground network; this is more pronounced when the offered voice load is high. Regarding the threshold and graceful degradation models considered for the correct reception of data packets, we have shown that, for low offered voice traffic loads, using the threshold model results in data throughput and delay performance which is better than obtained when the graceful degradation model is used, while the opposite is true for higher voice traffic loads. If such truncation were not necessary, the graceful degradation model would perform better than the threshold model for all cases of interest.

Overall, our modeling and performance analysis study of mixed-media packet radio networks with voice and data traffic can provide the tools for understanding these networks, as well as guidelines for the proper design of suitable multiple-access schemes.

Acknowledgement

The authors would like to express their sincere thanks to Dr. Jeffrey Wieselthier of the Naval Research Laboratory for the insightful advice and stimulating comments that he has offered throughout the preparation of this report.

REFERENCES

[1] D. Huynh, H. Kobayashi, and F. F. Kuo, "Optimal design of mixed media packet-switching networks: Routing and capacity assignment," *IEEE Trans. Communications,* vol. COM-25, January 1977.

[2] M. Schwartz, "Telecommunication networks: Protocols, Modeling and analysis," *Addison-Wesley,* 1987.

[3] F. A. Tobagi, "Multiaccess Protocols in Packet Communication Systems," *IEEE Transactions on Communications,* Vol. COM-28, April 1980.

[4] G. J. Coviello and P. A. Vena, "Integration of Circuit/Packet Switching by a SENET (Slotted Envelope Network) Concept," *1975 National Telecommunications Conference,* pp. 42.12-42.17.

[5] J. E. Wieselthier and A. Ephremides, "Fixed- and movable-boundary channel-access schemes for integrated voice/data wireless networks," network," submitted to the *IEEE Transactions on Communications,* May 1991.

[6] E. Geraniotis and M. Soroushnejad, "Performance evaluation of multi-access strategies for an integrated voice/data CDMA packet radio network," NRL Memorandum Report 6743, Naval Research Laboratory, Washington, DC, November 1990; also submitted to the *IEEE Transactions on Communications.*

[7] F. C. Schoute, "Dynamic frame length ALOHA," *IEEE Trans. Communications,* vol. COM-31, April 1983.

[8] J. E. Wieselthier, A. Ephremides, and L. A. Michaels, "An Exact Analysis and performance evaluation of framed ALOHA with Capture," *IEEE Trans. Communications,* vol. COM-37, February 1989.

[9] W. Feller, *An Introduction to Probability Theory and its Applications,* vol. I, 3rd ed. New York:Wiley, 1968, p. 112.

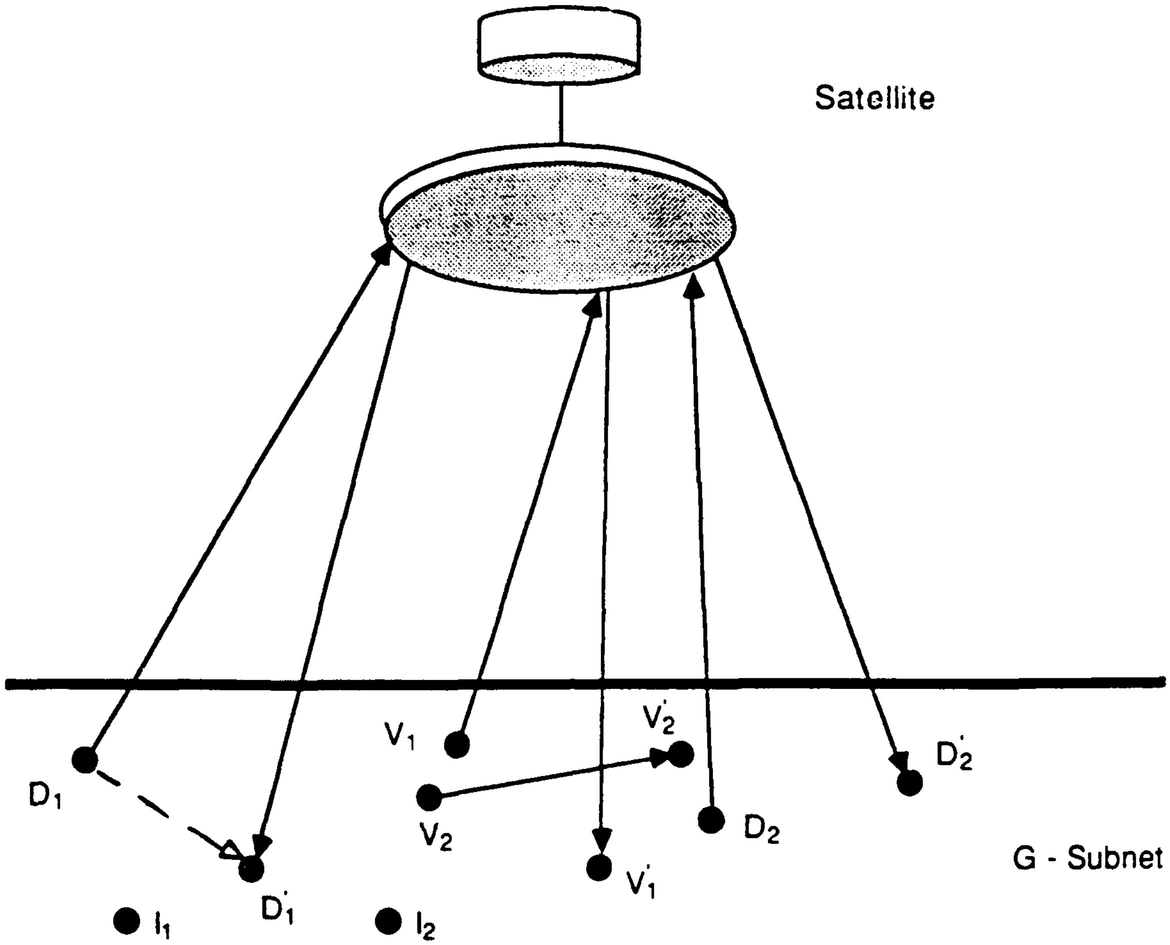
[10] B. Hajek and T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automatic Control,* Vol. AC-27, pp. 559-569, June 1982.

[11] B. Hajek, "Recursive retransmission control-Application to a frequency-hopped spread-spectrum system," *Proceedings of 1982 Conference on Information and Systems Sciences,* Princeton, New Jersey, March 1982.

[12] E. Geraniotis. "The probability of multiple correct packet receptions in a multi-receiver frequency-hopped spread-spectrum system." NRL Memorandum Report 6882, Naval Research Laboratory, Washington, DC, September 1991. A version of this report entitled "Multireception probabilities for FH/SSMA communications" will be published in the *IEEE Transactions on Communications* in Spring 1992.

[13] E. Geraniotis and M. B. Pursley, "Error Probabilities for Slow-Frequency-Hopped Spread-Spectrum Multiple-Access Communications over Fading Channels." Special Issue on Spread-Spectrum Communications of the *IEEE Transactions on Communications,* Vol. COM-30, May 1982.

[14] D. Bertsekas and R. Gallager, *Data Networks.* Prentice Hall, Engelwood Cliffs, New Jersey, 1987.



D: Active Data Nodes
V: Active Voice Nodes
I: Idle Nodes
——— Links with Direct Transmissions or Retransmissions
- - - - - Links with Retransmissions only

Figure 1. Mixed-Media (Satellite/Ground) Network

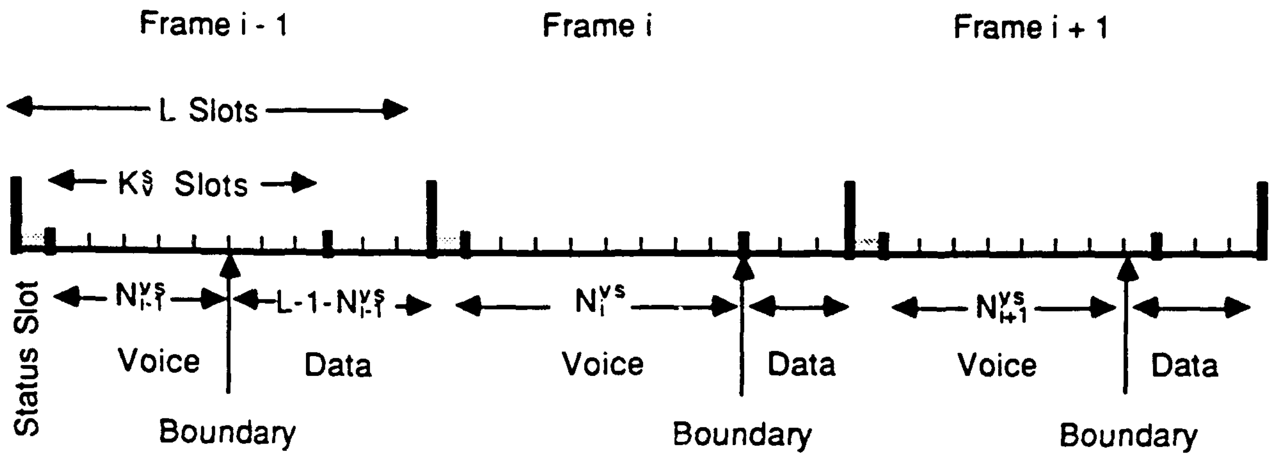


Figure 2. Satellite Subnetwork Using Framed ALOHA

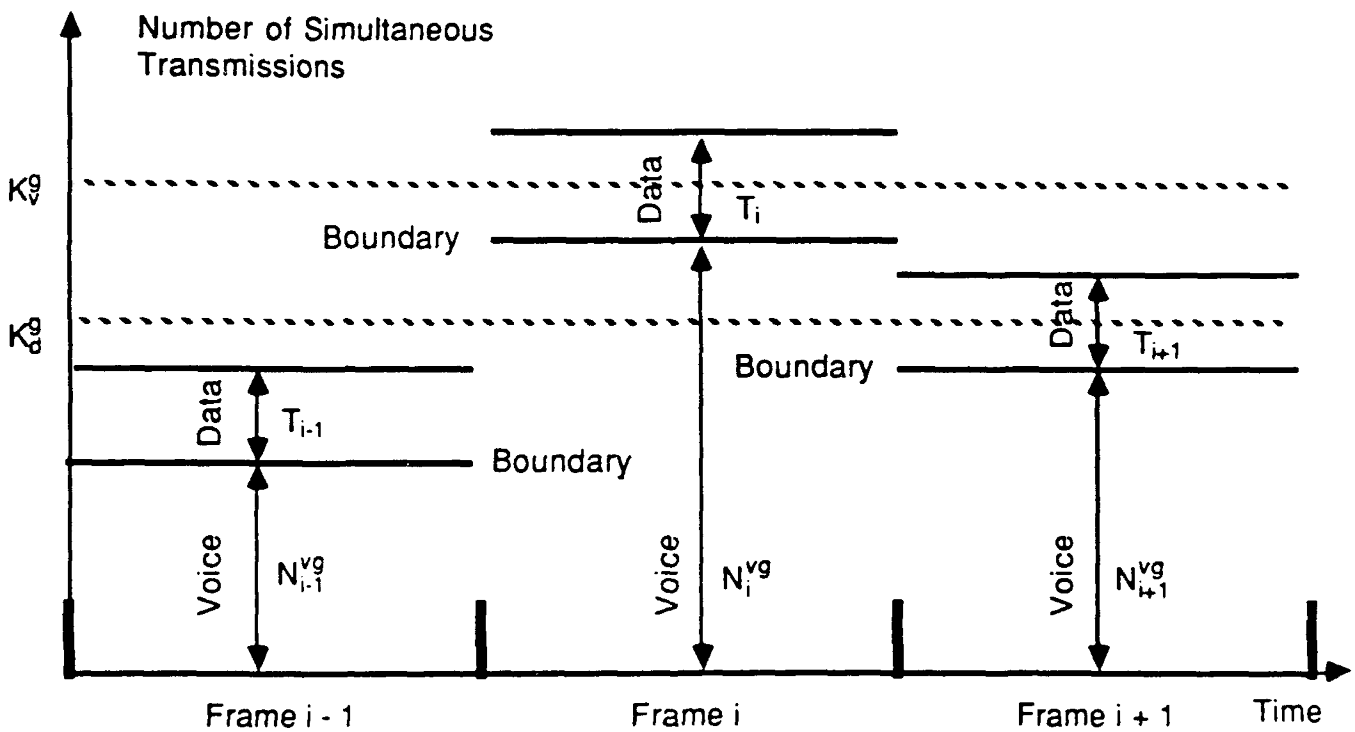


Figure 3. Ground Subnetwork Using CDMA

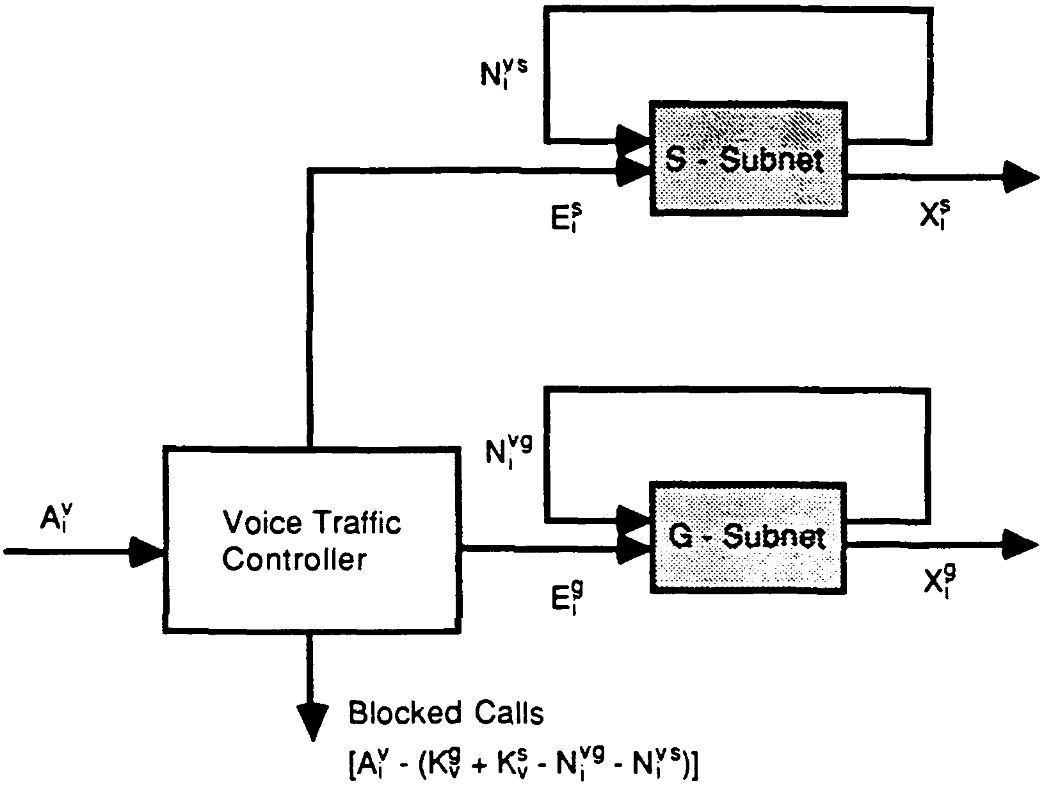


Figure 4. Voice Traffic over Ground and Satellite Subnets

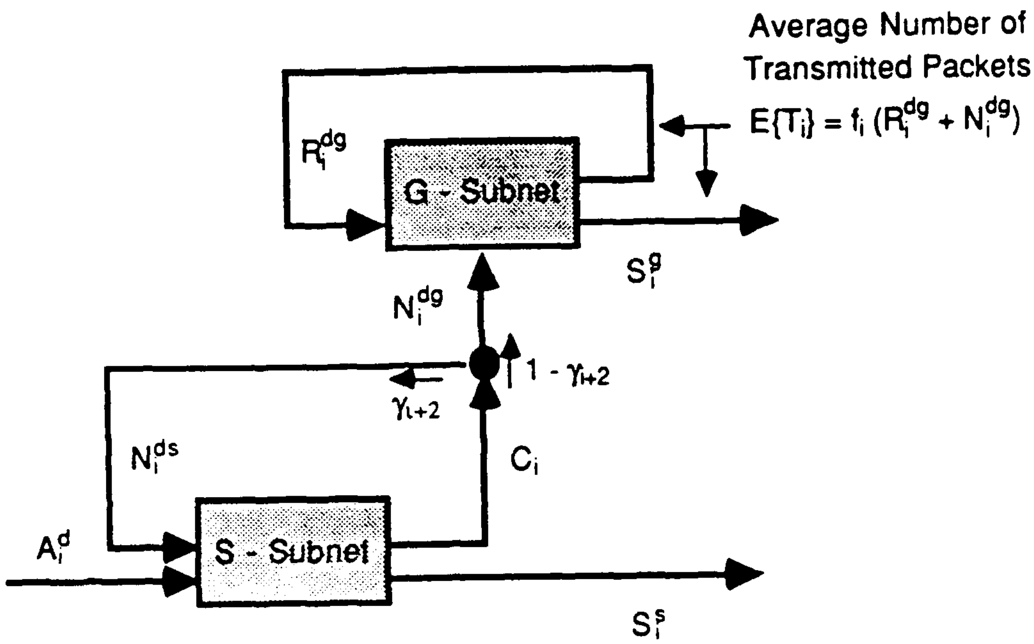


Figure 5. Data Traffic over Satellite and Ground Subnets

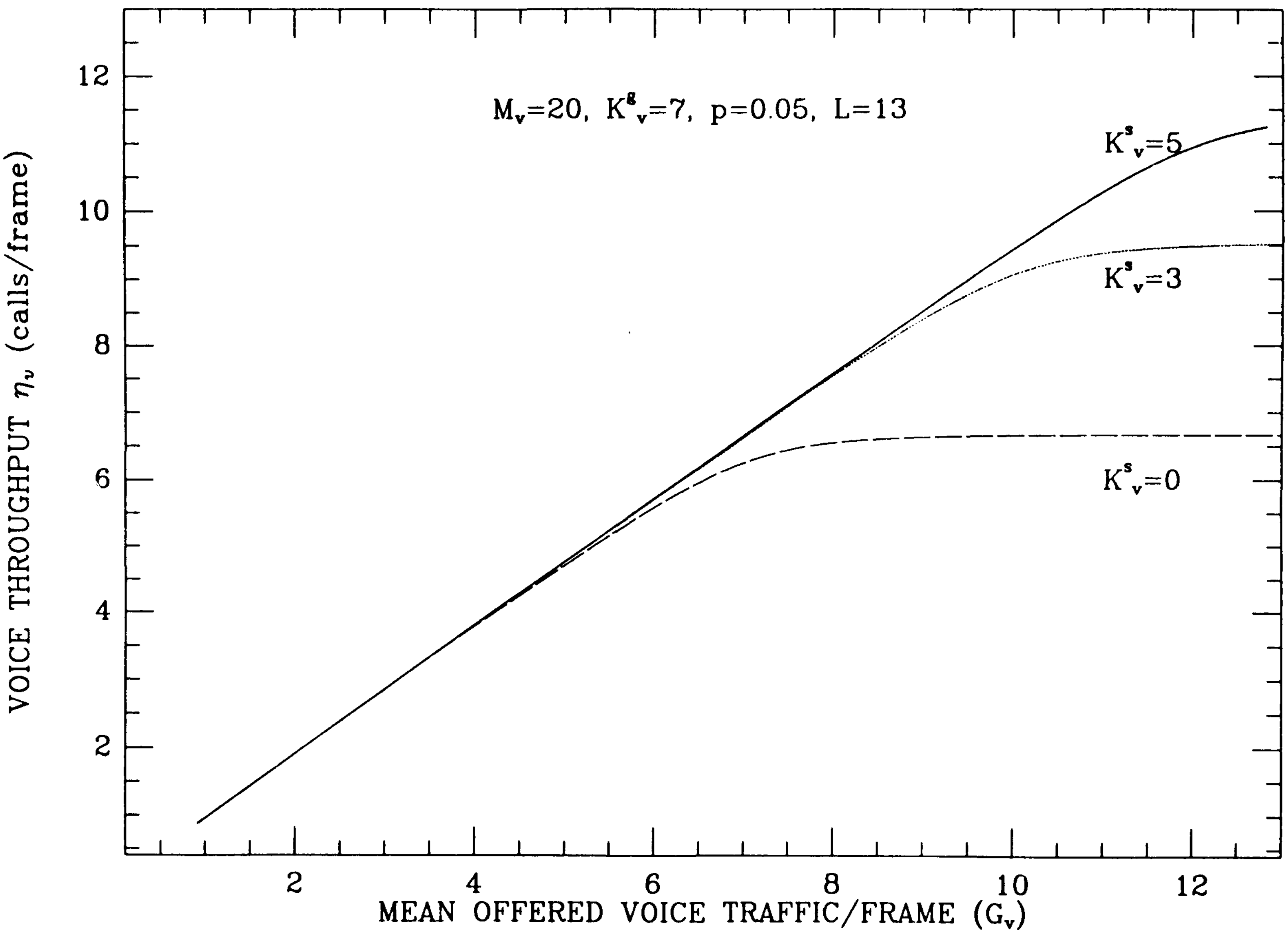


Figure 6 Voice throughput over S– and G– nets

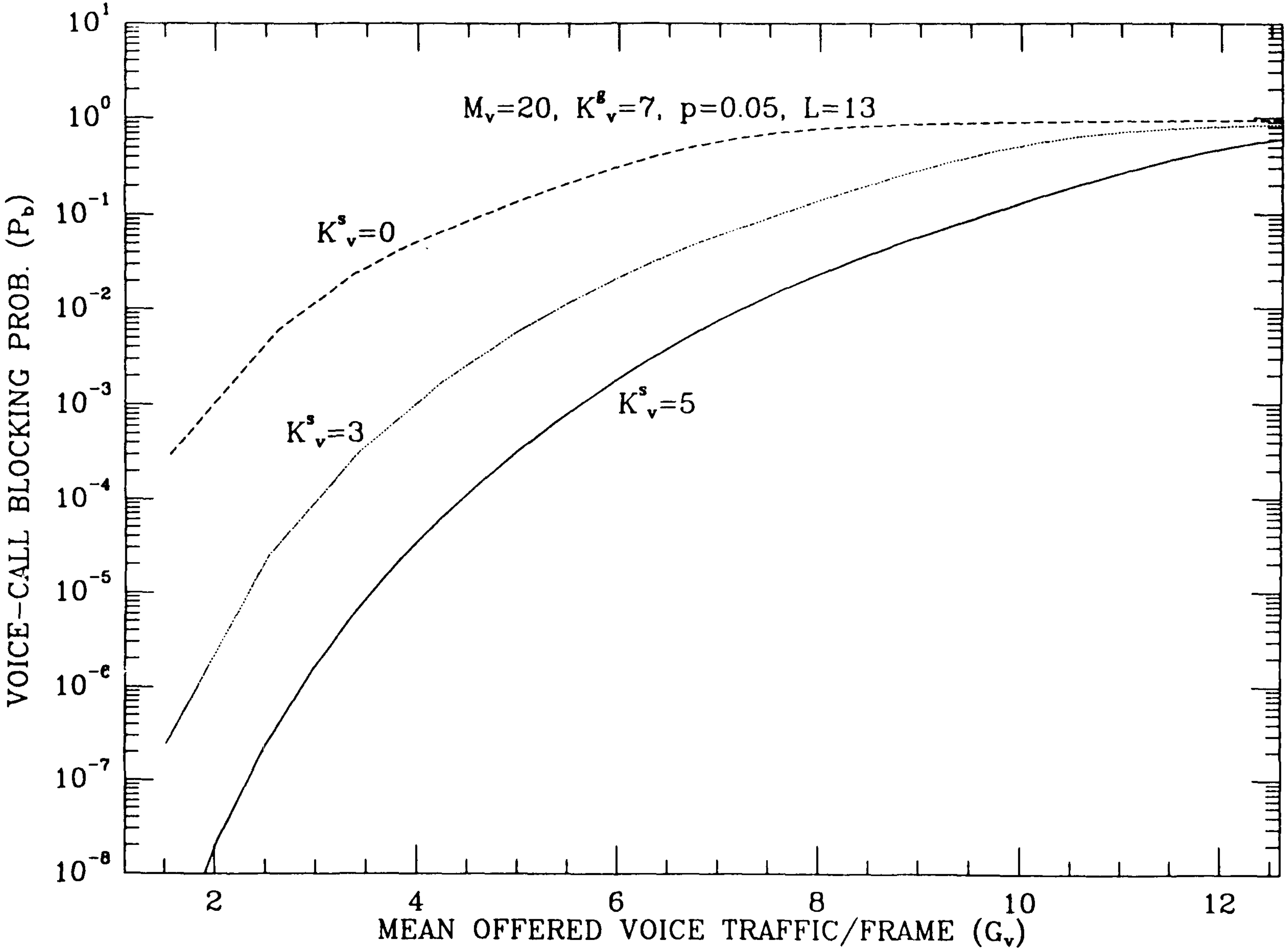


Figure 7 Voice blocking probability over S- and G- nets

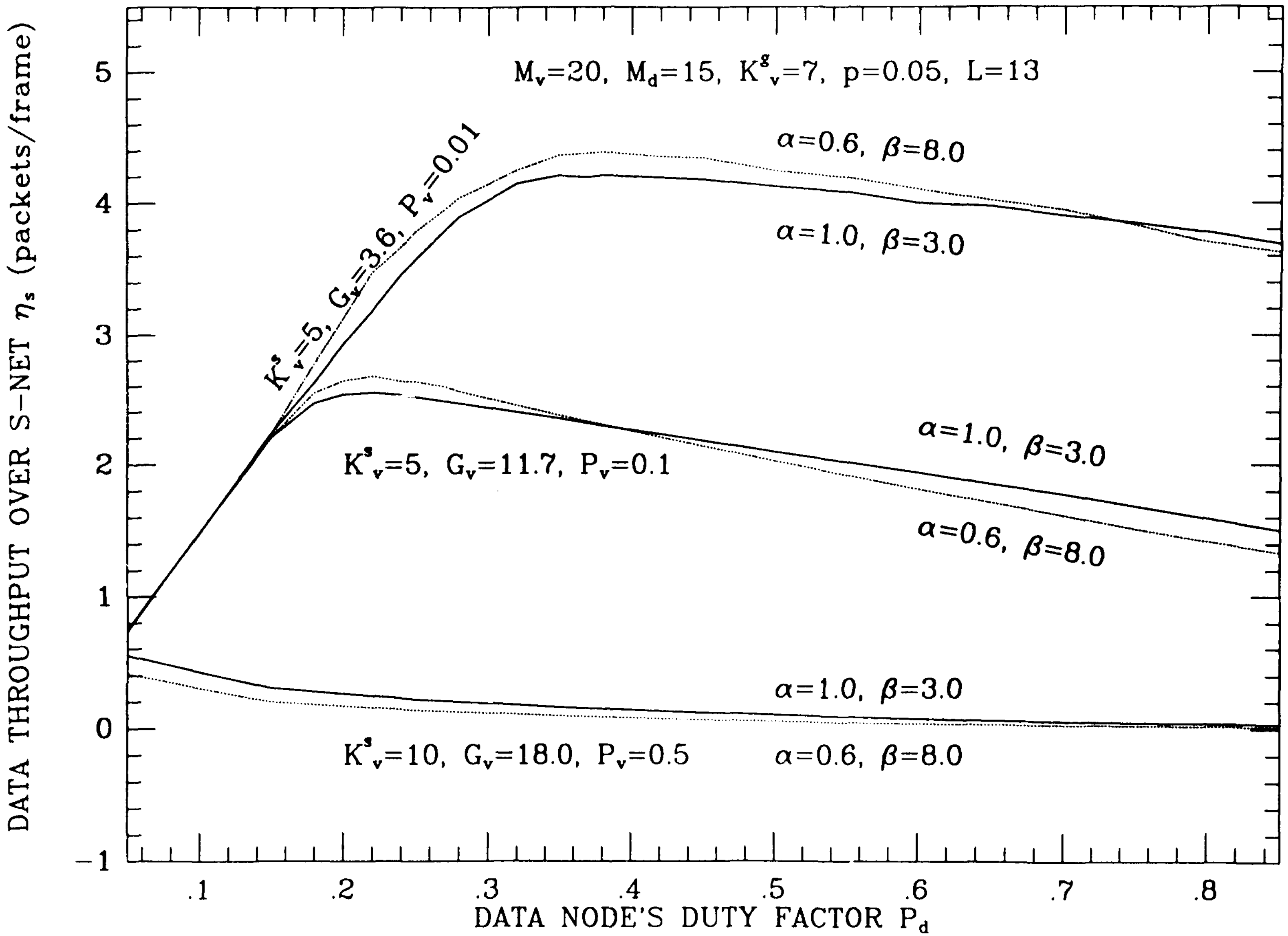


Figure 8 Data throughput over S-net

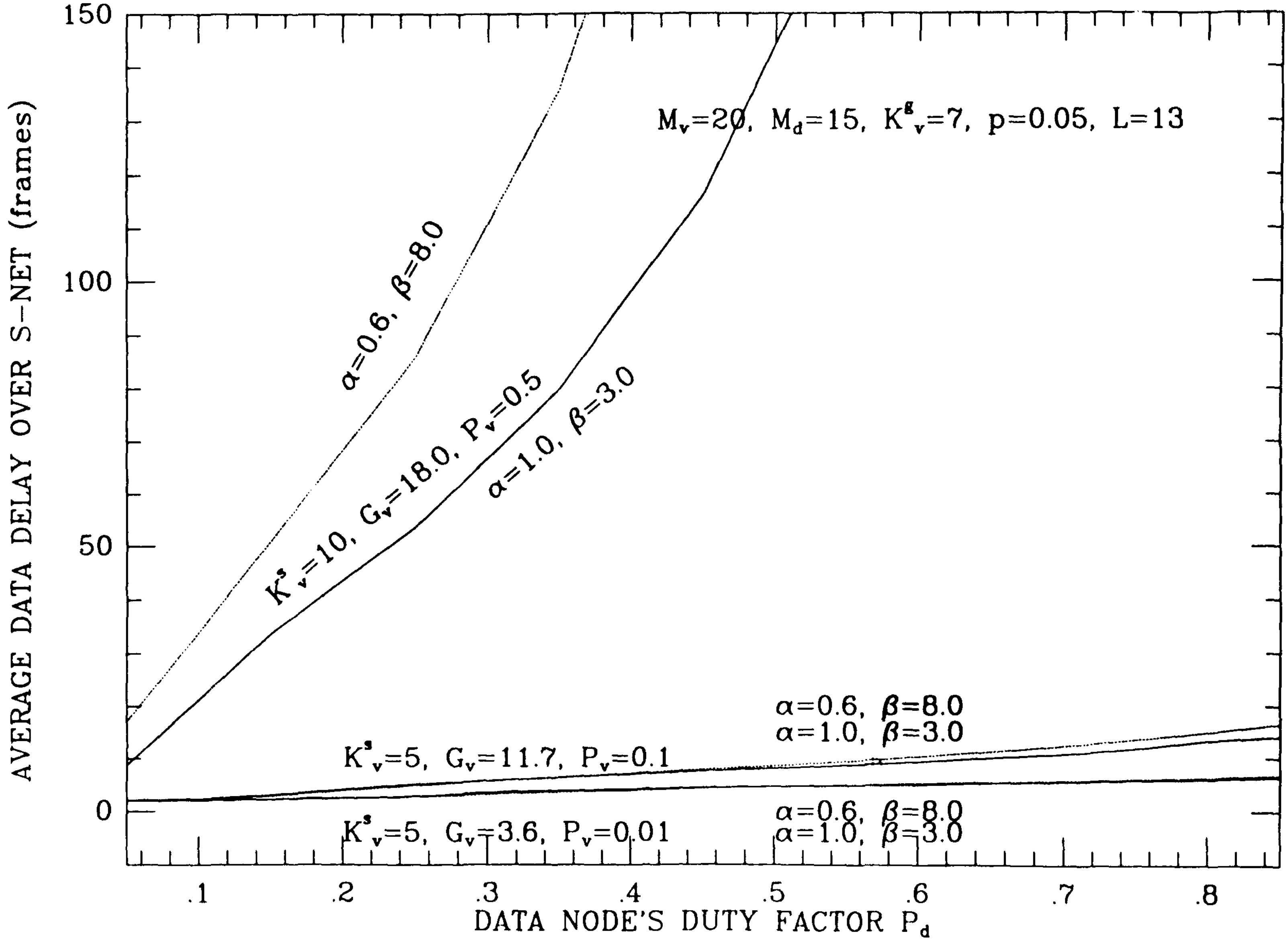


Figure 9 Average data delay over S-net

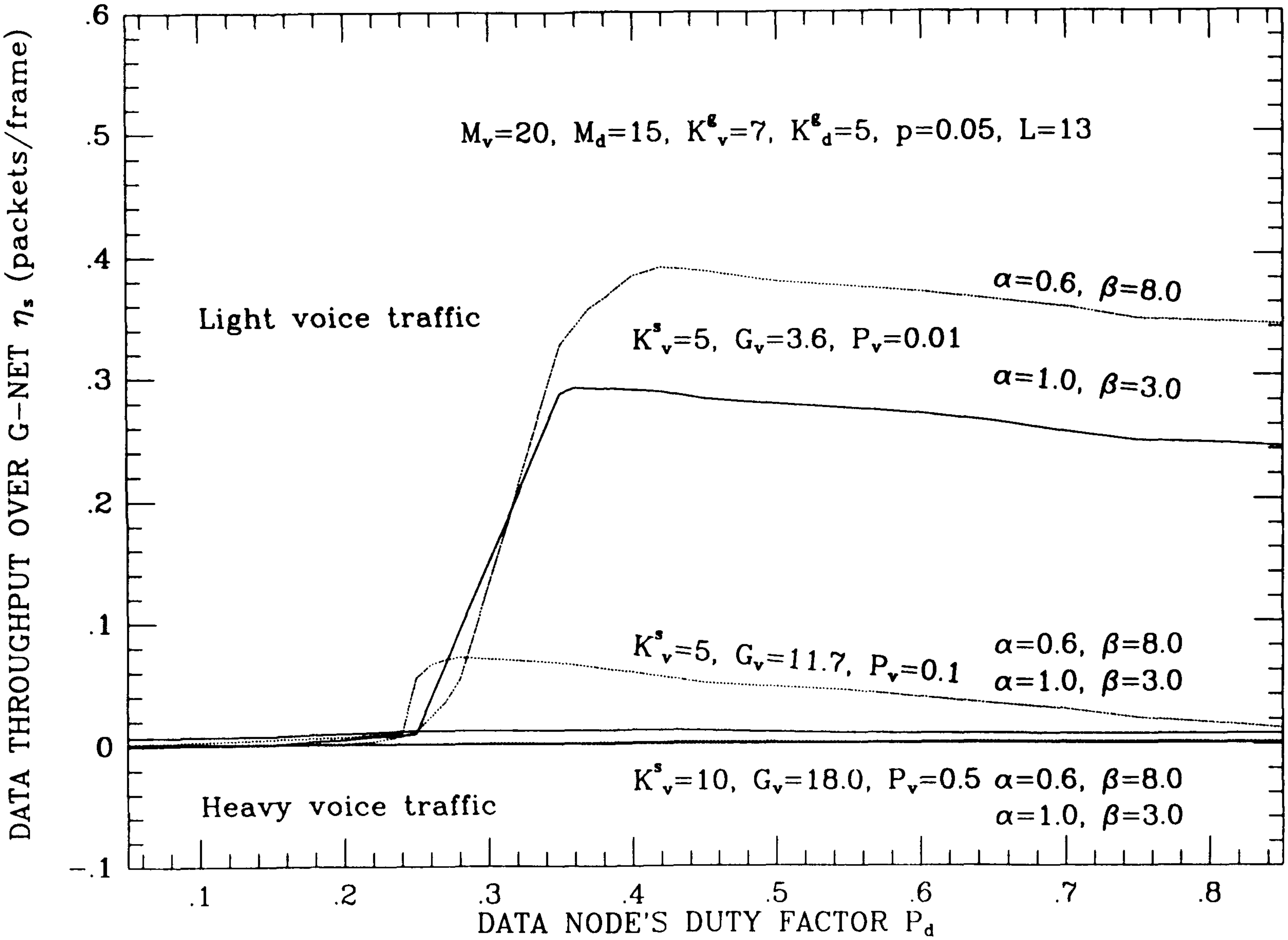


Figure 10 Data throughput over G-net (Threshold Model)

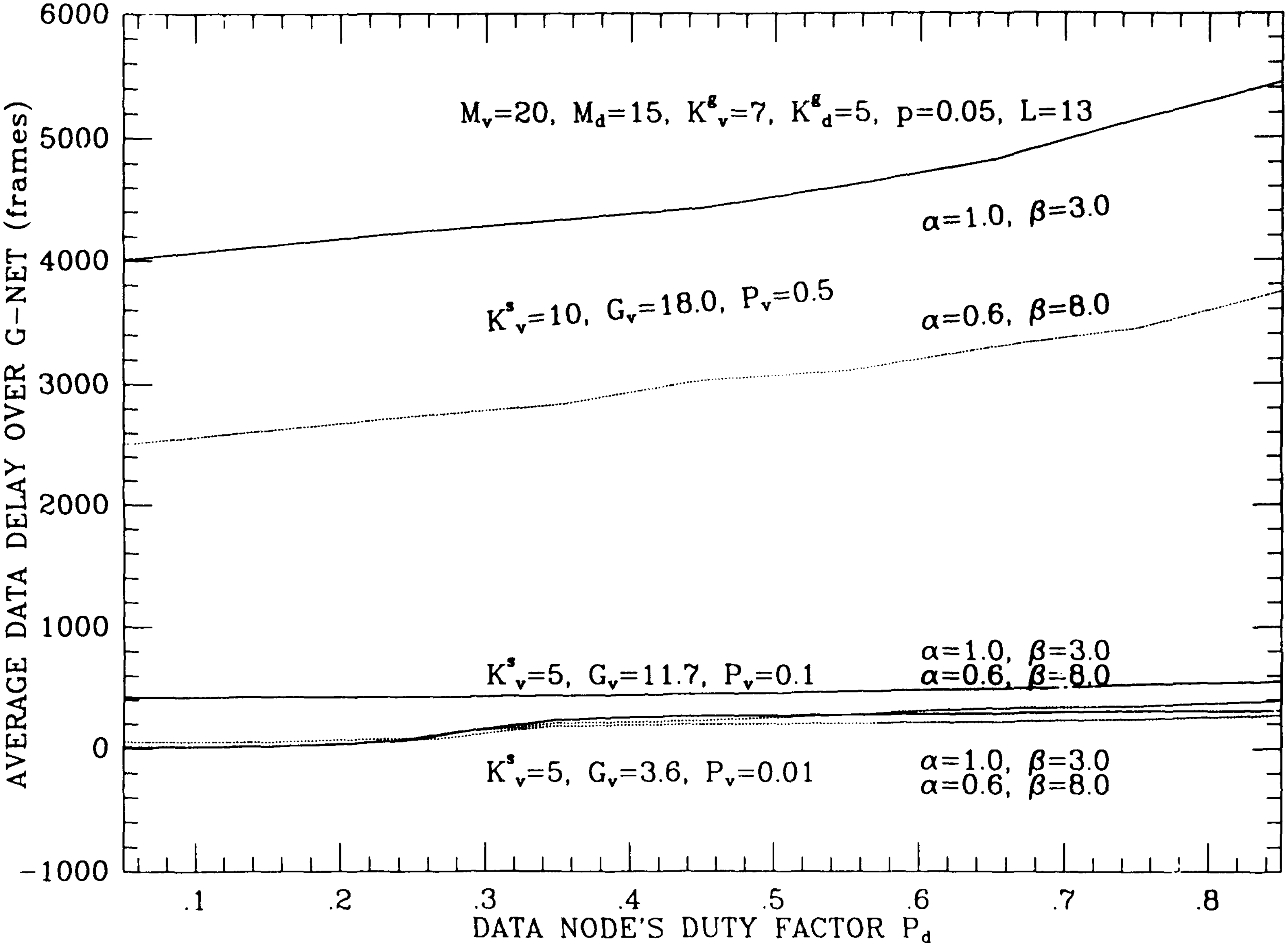


Figure 11 Data delay over G−net (Threshold Model)

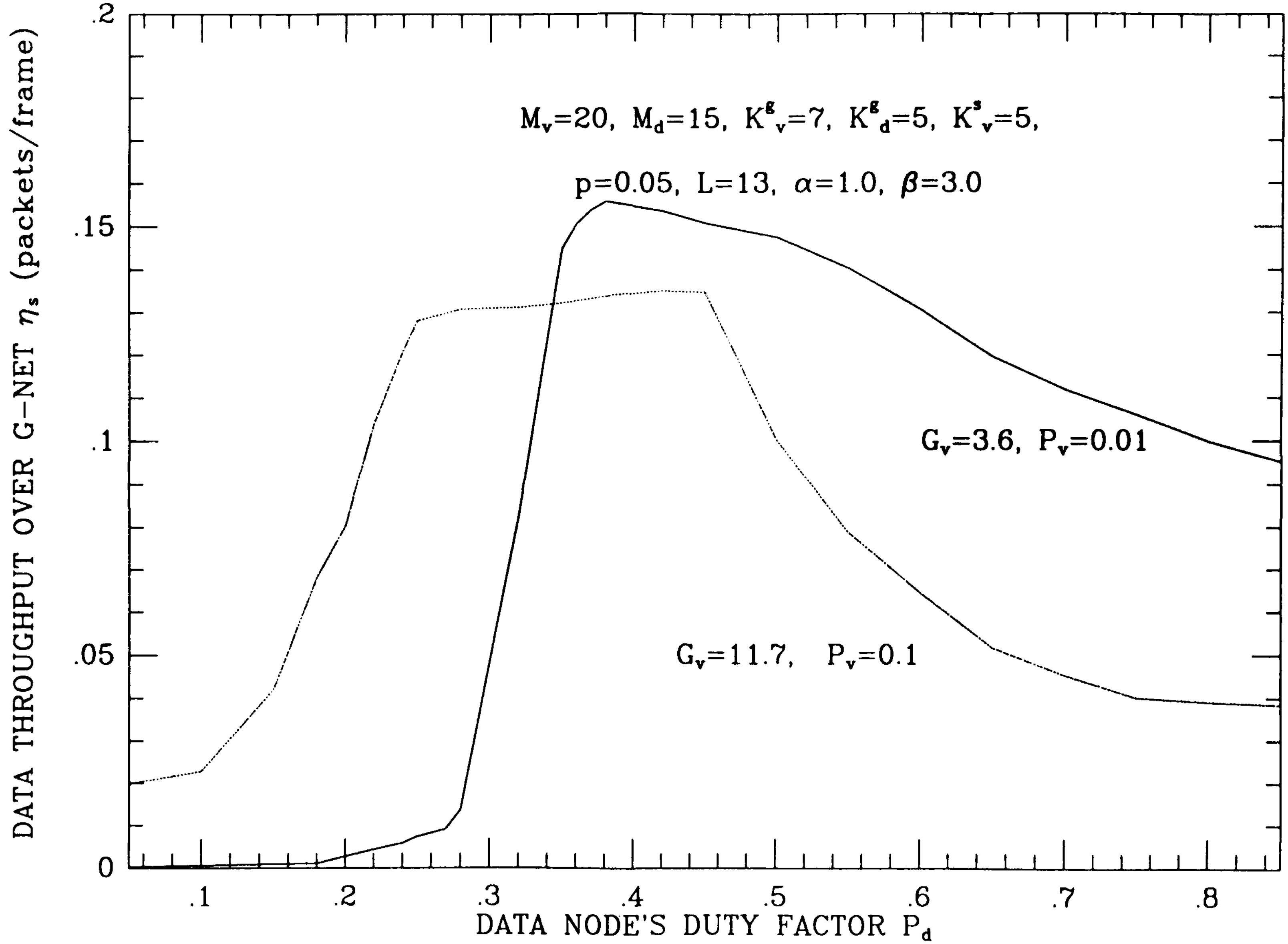


Figure 12 Data throughput over G-net (Graceful degradation model)

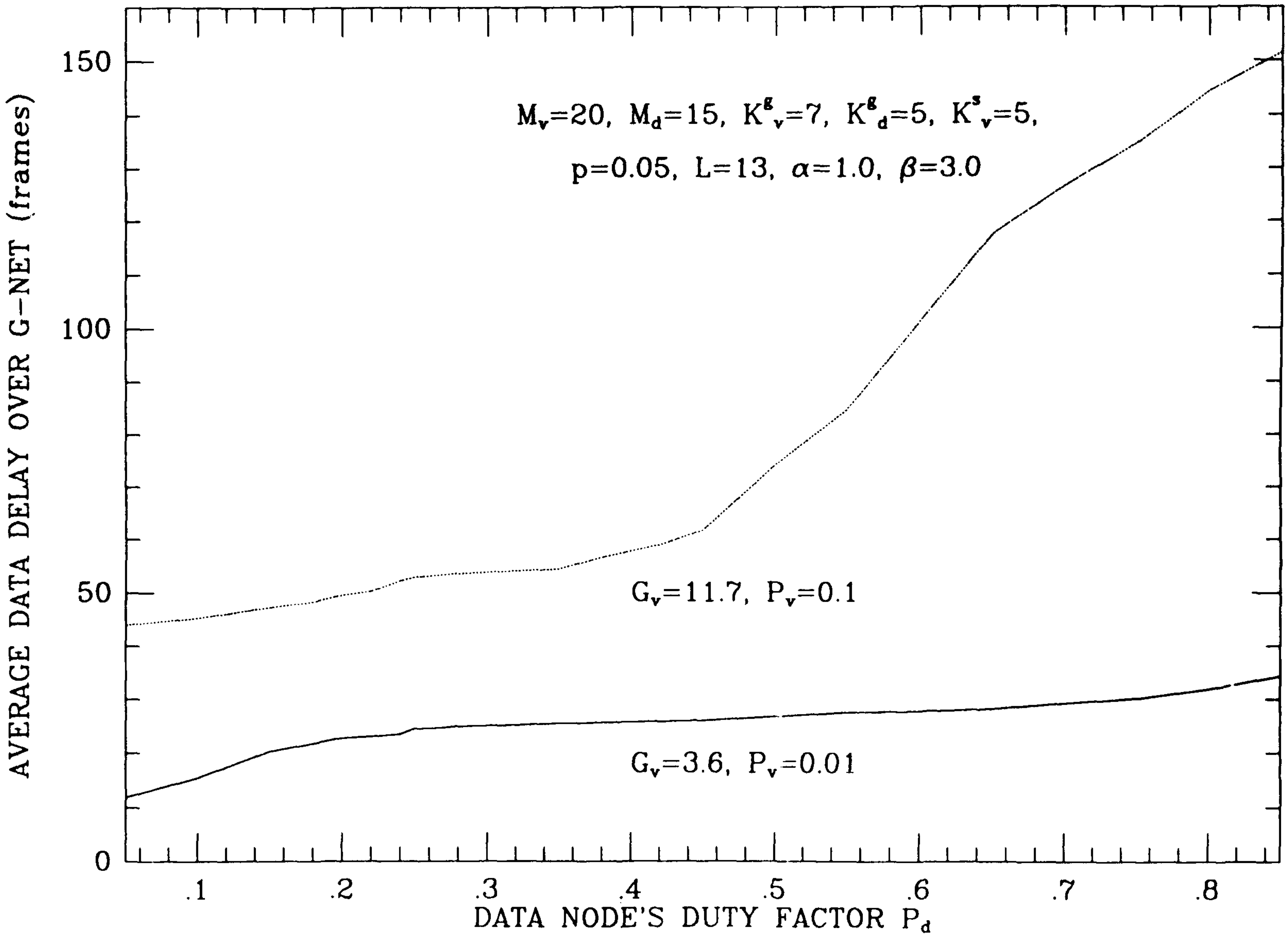


Figure 13 Data delay over G-net (Graceful degradation model)

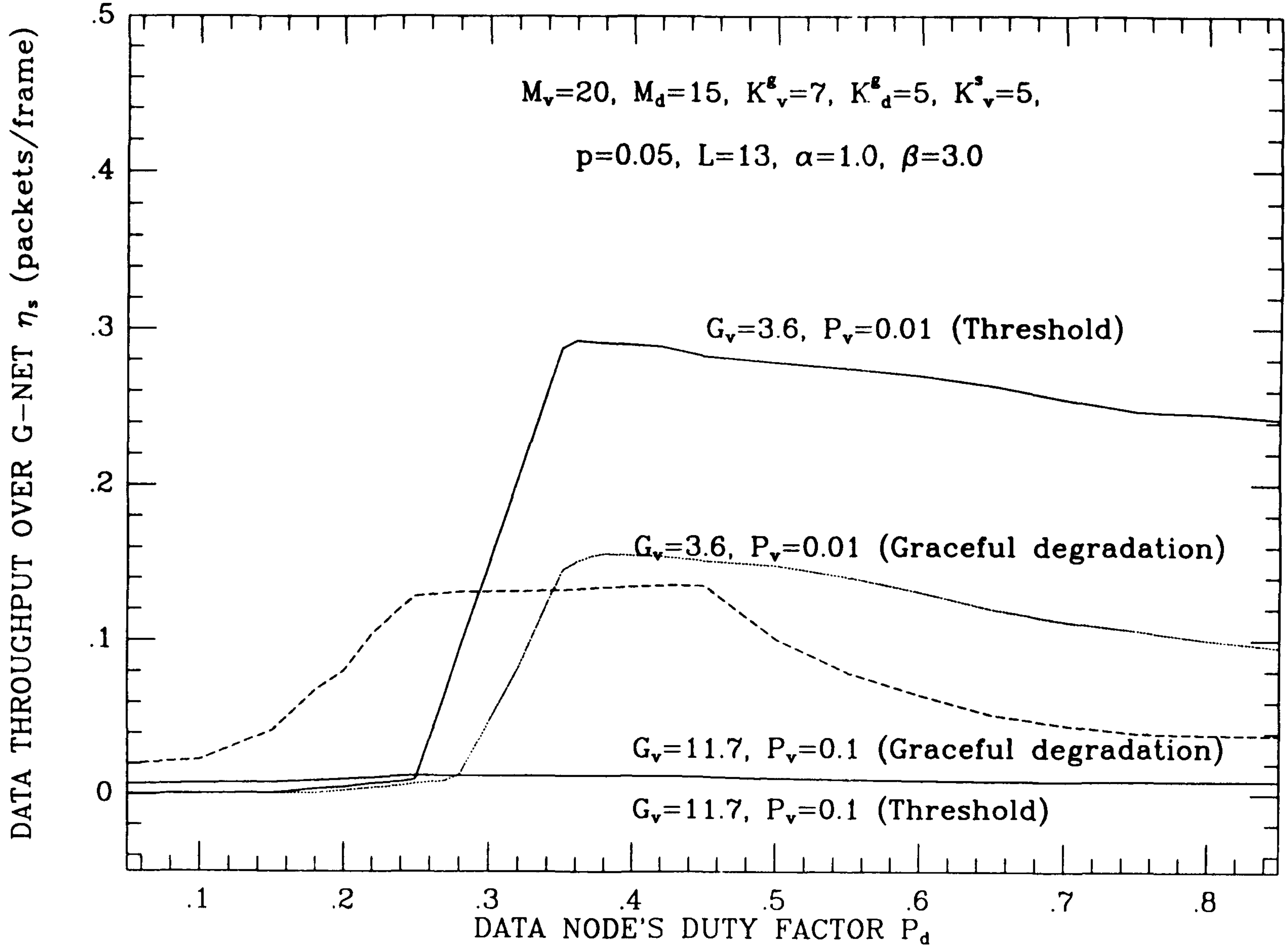


Figure 14 Data throughput over G-net for the threshold and graceful degradation models

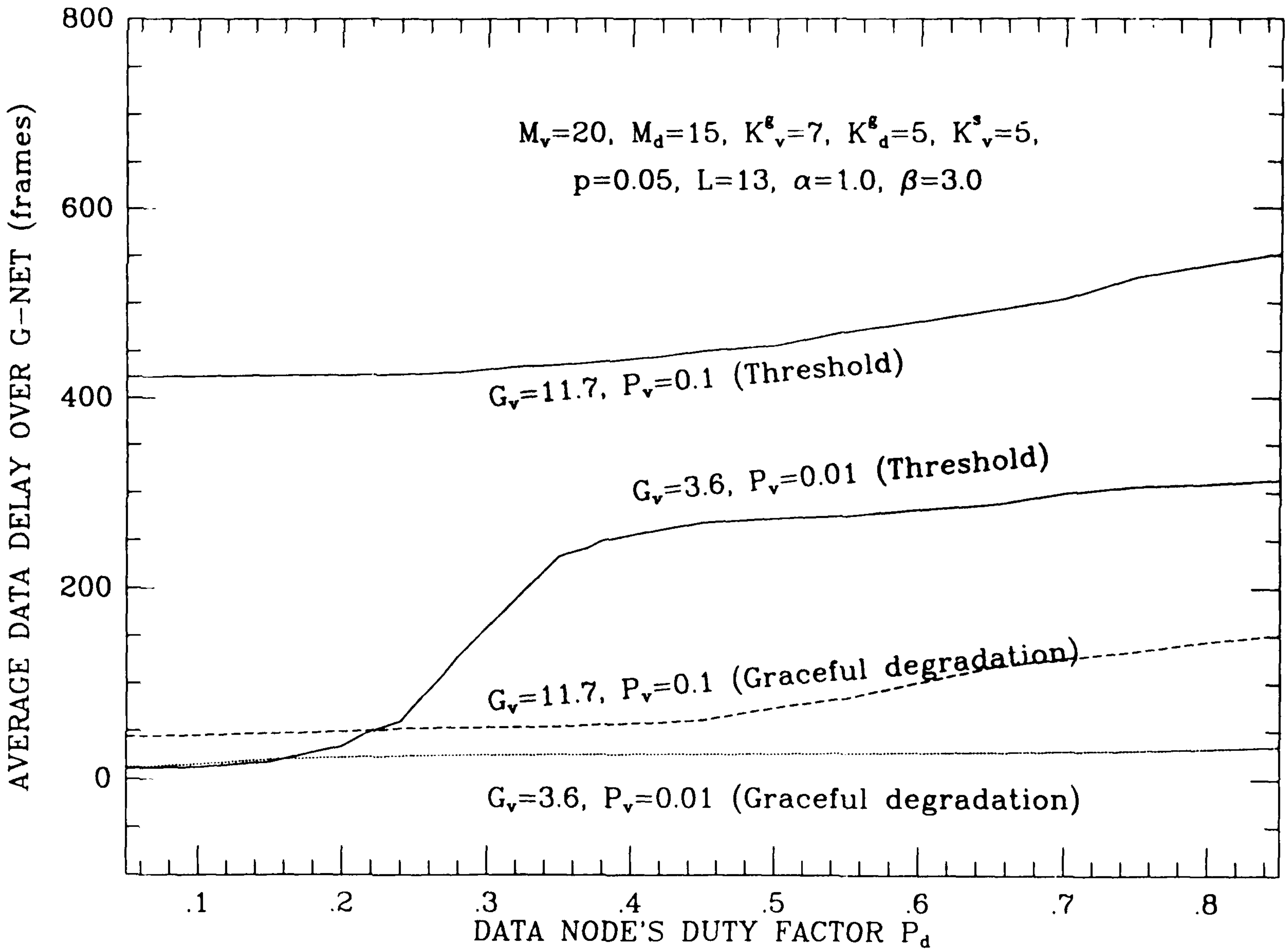


Figure 15 Data delay over G-net for the threshold and graceful degradation models

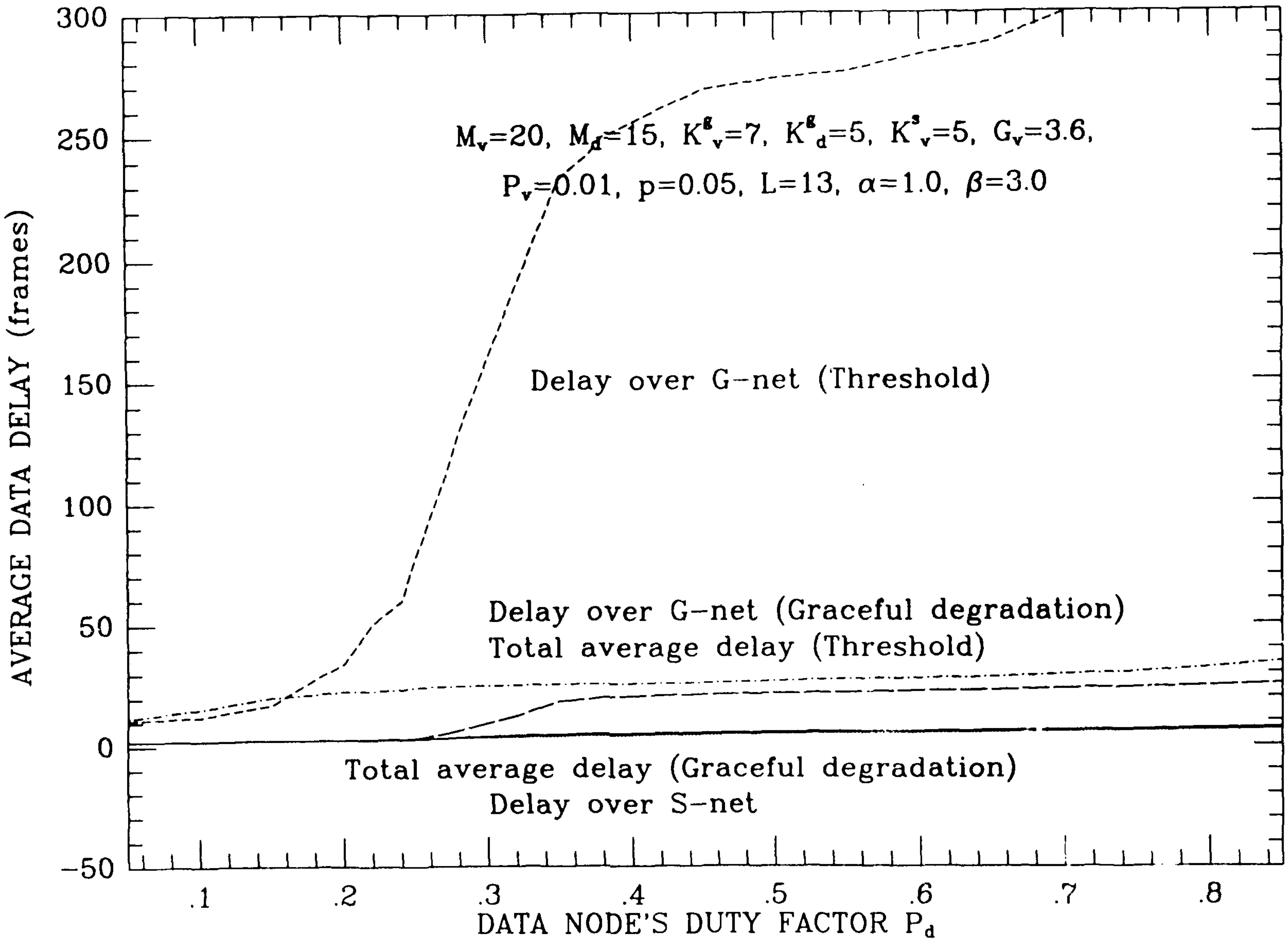


Figure 16 Data delays for light voice traffic

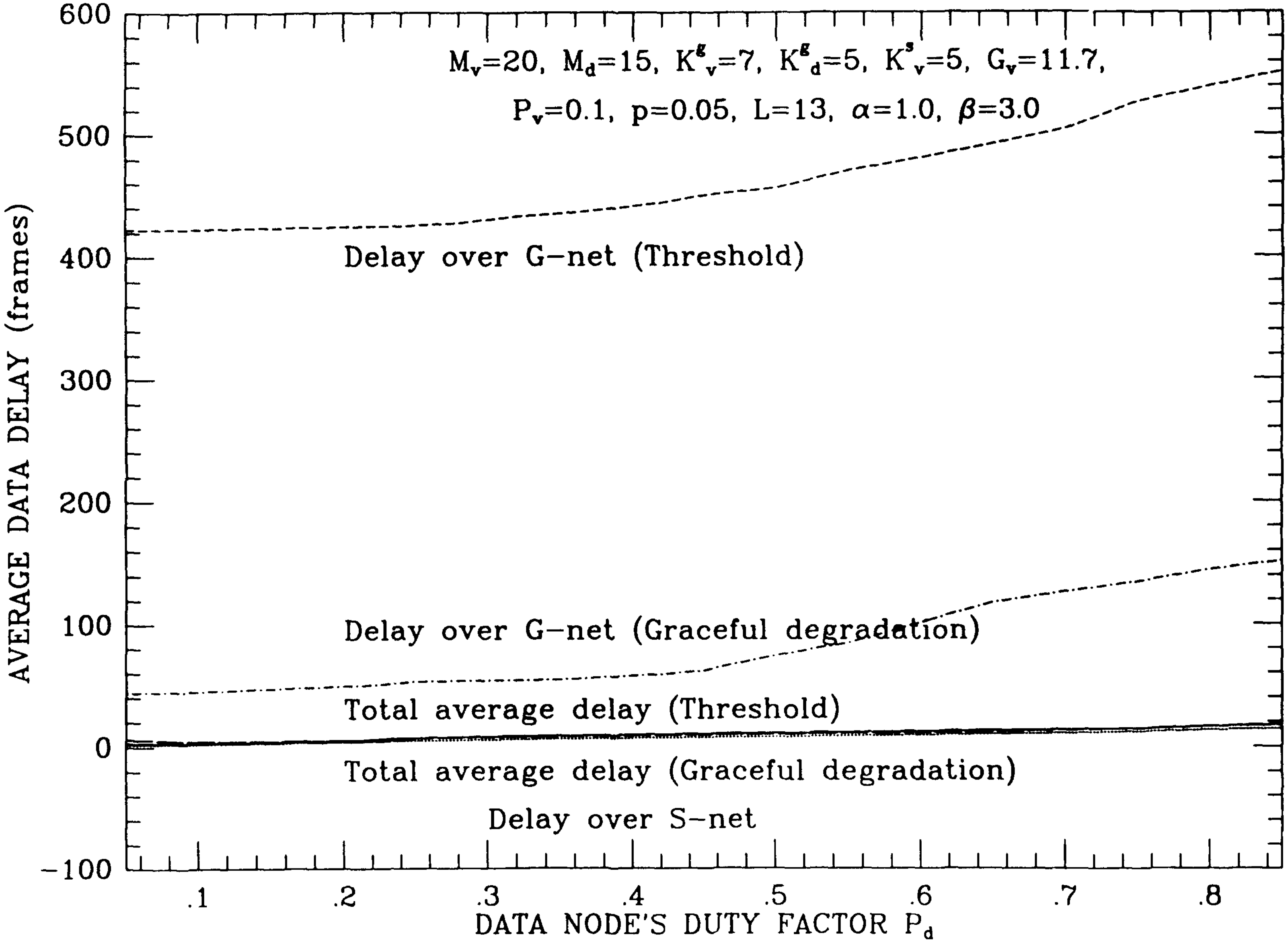


Figure 17 Data delays for heavy voice traffic